



Universidade do Porto

FEUP Faculdade de
Engenharia

Dissertação submetida para a satisfação parcial
dos requisitos do grau de mestre em Estatística Aplicada e Modelação

Escolha de Variáveis

em Análises Estatísticas Multidimensionais

Maria Isabel Pinto Preto

Dissertação realizada sob supervisão do
Professor Doutor Francisco José Lage Campelo Calheiros
do Departamento de Engenharia Civil, Secção de Matemática e Física
da Faculdade de Engenharia da Universidade do Porto

Mestrado em Estatística Aplicada e Modelação
Porto, 2005

‘Quando se encontra sem procurar é porque muito se procurou sem encontrar’

Popular

Agradecimentos

À Faculdade de Engenharia da Universidade do Porto pelas condições proporcionadas e ao Mestrado em Estatística Aplicada e Modelação, em particular à sua directora, Professora Teresa Arede.

Ao meu orientador, Professor Francisco Calheiros. Pelo engenho. Pela total disponibilidade. Pela preocupação constante. E sobretudo...pela imensa paciência cheia de boa disposição!

Aos meus pais... Por tudo. Pela oportunidade, por sempre acreditarem e pela segurança de saber que estão sempre presentes ...

E porque há muito formas de contribuir:

Ao meu irmão Chico...

À D. Aida porque nunca deixa de se preocupar e torcer para que tudo corra bem! E claro, por aquela paciência...

Aos amigos de sempre: Soraia, Vera e Rui! Porque a amizade é muitas vezes o maior contributo!

Ao primo Bruno pela companhia constante e pela disponibilidade!

À prima Marta pelas conversas à distância, tantas vezes fonte de inspiração e motivação!

Resumo

Num estudo estatístico, o uso de variáveis tal como foram observadas pode ser pobre e, eventualmente, não descrever aspectos importantes.

A existência de estruturas *a priori* nos dados pode sugerir a construção de novas variáveis. Em situações puramente geométricas, das variáveis medidas (comprimentos) podem-se construir volumes, áreas, medidas de forma (cubicidade, por exemplo), etc. Em situações ligadas ao tempo, a estrutura temporal pode induzir à construção de tendências e sazonalidades.

Um objectivo de uma análise estatística é utilizar o melhor possível a informação disponível. Será que variáveis construídas a partir das variáveis observadas permitem melhorar o uso da informação? Será que se pode ou deve aumentar o número de variáveis indiscriminadamente, quase sem limite?

Evidentemente, que a resposta à última pergunta é negativa. Porém, é conveniente esclarecer quantas e que variáveis se devem escolher em cada caso.

Analisamos três tipos de exemplos:

- um caso puramente geométrico em que os objectos em estudo são paralelepípedos;
- notas de alunos, entre 1 e 5, que tentamos assimilar a paralelepípedos;
- consumos bimensais das Águas de Gaia.

Muitas vezes é a organização das variáveis em grupos, conjuntamente, com o *número* de variáveis de cada grupo que fixa as componentes principais.

A organização das variáveis em grupos com correspondentes blocos na matriz de correlações levou-nos à análise experimental destas matrizes permitindo enunciar uma nova conjectura, que estabelece uma relação entre o *número* de variáveis em cada grupo e os valores próprios.

Abstract

In a statistical study, the use of variables such as they had been observed can be poor and, eventually, not to describe important aspects.

The existence of *a priori* structures in the data can suggest the construction of new variables. For example, in purely geometric situations, from measured variables (lengths) can be constructed volumes, areas, form measures (cubicity), etc. In situations involving time, the temporal structure can induce to the construction of trends and seasonalities.

The objective of a statistics analysis is to use in the best possible way the available information. Will constructed variables from the original ones allow to improve the use of the information? Is it possible to increase indiscriminately the number of used variables, almost without limit?

The answer to the last question is, clearly, negative. However, it's convenient to clarify how many and what variables should be chosen, in each case.

We analyze three types of examples:

- a purely geometric case where the objects in study are parallelepipeds;
- notes of pupils, between 1 and 5, that we try to assimilate to parallelepipeds
- consumptions in two consecutive months of Águas de Gaia.

Many times is the variables organization in groups jointly with the *number* of variables in each group that fixes the principal components.

The organization of the variables in groups with corresponding blocks in the correlation matrix led us to its experimental analysis which allowed to enunciate a new conjecture, that establishes a relation between the *number* of variables in each group and the *eigenvalues*.

Organização da Tese

Agradecimentos	i
Resumo.....	ii
Abstract	iii
Organização da Tese.....	iv
Capítulo I – Introdução	1
Capítulo II – Correlações com outros Trabalhos	4
Capítulo III – Escolha de Variáveis num Contexto Geométrico	7
3.1 O exemplo de J. Davis.....	7
3.2 Paralelepípedos Rectângulos	9
3.3 Estatística Descritiva das Variáveis	11
3.4 Análise em Componentes Principais	13
3.4.1 Matriz de Correlações.....	13
3.4.2 Valores Próprios da Matriz de Correlações	14
3.4.3 Interpretação das Componentes Principais	15
3.4.3.1 Interpretação da Primeira Componente Principal	16
3.4.3.2 Interpretação da Segunda Componente Principal	19
3.4.3.3 Interpretação da Terceira Componente Principal.....	21
3.5 Estabilidade da ACP realizada.....	24
3.5.1 Variação do primeiro valor próprio em função da dimensão	24
3.5.2 Variações Angulares.....	30
3.5.2.1 Primeira Componente	30
3.5.2.2 Segunda Componente	31
3.5.2.3 Terceira Componente.....	33
3.5.2.4 Análise da Normalidade do Ângulo	34
3.6 Comparação com Indivíduos FC-FS	36
Capítulo IV – Variantes	38
4.1 Paralelepípedos de Dimensões não ordenadas.....	38
4.2 Hiper-Paralelepípedos de Dimensão Quatro.....	43
4.3 Caso Contínuo.....	55

4.3.1 Estatística Descritiva das Variáveis	55
4.3.2 Análise em Componentes Principais.....	57
4.3.2.1 Matriz de Correlações.....	57
4.3.2.2 Valores Próprios da Matriz de Correlações	58
4.3.2.3 Interpretação das Componentes Principais	58
4.4.2.3.1 Interpretação da Primeira Componente Principal ..	59
4.4.2.3.2 Interpretação da Segunda Componente Principal ..	61
4.4.2.3.3 Interpretação da Terceira Componente Principal ..	62
4.3.3 Estabilidade da ACP realizada.....	65
4.3.3.1 Variação do primeiro valor próprio em função da dimensão..	65
4.3.3.2. Variações Angulares.....	69
4.4 Classificações de alunos	73
4.5 Caso Multinormal.....	88
Capítulo V – Variáveis FC-FS	90
Capítulo VI – Matrizes Organizadas em Blocos.....	96
6.1 Matrizes organizadas em dois blocos	96
6.2 Variação dos Maiores Valores Próprios em Função de n , δ e ε	101
6.3 Organização do Círculo de Correlações	106
Capítulo VII – Conclusões	111
Bibliografia.....	115
Anexos	116
Anexo I – Perfis de Consumidores	117
Anexo II – Documentação do Programa	124
Anexo III – Paralelepípedos de J. Davis	129
Anexo IV – Paralelepípedos Inteiros	130
Anexo V – Hiper-Paralelepípedos de Dimensão 4	135
Anexo VI – Paralelepípedos – Caso Contínuo	140
Anexo VII – Classificações de Alunos	144

Capítulo I

Introdução

Esta tese tem como tema a escolha de variáveis antes de se proceder à análise dos dados.

Dado um conjunto de dados como tratá-lo? Que variáveis escolher e quais as que são úteis à sua interpretação?

Foi-nos fornecido pelas Águas de Gaia (Professor Doutor Poças Martins) um ficheiro contendo os consumos bimensais de utilizadores da rede pública de água, referentes ao período compreendido entre Janeiro de 1999 e Fevereiro de 2003, num total de 10090251 registos, com o objectivo de organizar os perfis de consumidores (Anexo I).

Os problemas fundamentais encontrados no seu estudo foram a organização da base de dados (eliminação de dados aberrantes, tentativa de completamento de dados em falta, etc.), escolha e definição de variáveis que sintetizassem, em algum sentido, as características dos consumidores.

As alternativas passavam por considerar uma simples classificação por tamanhos ou decidir pela procura de padrões de consumo. Perante estas opções, decidiu-se construir, a partir dos dados em bruto dos consumos bimensais, diferentes variáveis que permitissem contrastar distintos aspectos. As variáveis escolhidas incluíram tendências, médias e extremos.

Não tendo sido um sucesso completo, subsistiu a dúvida se a escolha das variáveis terá sido a mais adequada para a definição dos perfis dos consumidores e qual o critério para avaliar a adequação.

Por isso, no capítulo II, retomou-se um exemplo geométrico, parcialmente considerado em FC-FS ([1]), proveniente de Davis ([2]) e Cooley & Lohnes ([3]): paralelepípedos definidos pelas dimensões das três arestas (ordenadas). Construíram-se paralelepípedos com arestas de comprimentos inteiros, variando entre um e dez (inclusive) e construíram-se variáveis função das medidas das arestas.

As escolhas efectuadas em termos de variáveis foram totalmente *ah-doc/experimentais* e sem se avaliar a hipótese de algumas das variáveis se poderem revelar de maior ou menor importância para a análise em componentes principais (ACP) realizada.

Foram empregues variáveis analíticas (polinómios em várias variáveis), ligadas a extremos (máximos e mínimos) e diferentes combinações de potências, como por exemplo, quocientes.

Dado o exemplo ser geométrico, todas as variáveis têm fácil interpretação.

Os exemplos dos paralelepípedos e das Águas de Gaia são especiais uma vez que existe uma estrutura interna dos dados. Em ambos os casos são efectuadas medições de variáveis do mesmo tipo: as três dimensões para os paralelepípedos e, nas Águas de Gaia, os consumos bimensais ao longo de 4 anos.

No caso das Águas de Gaia existe uma estrutura temporal subjacente, não sendo conveniente a ordenação dos consumos dado perder-se a referida estrutura. Considerando esta circunstância, no capítulo IV, procedeu-se ao estudo da variação da organização dos paralelepípedos não ordenando as respectivas dimensões.

No capítulo IV são também realizadas experiências com outros tipos de variáveis. Investiga-se o efeito da discretização considerando um caso contínuo (paralelepípedos cujas dimensões das arestas seguem uma distribuição uniforme $]0,10]$) e estudam-se hiper-paralelepípedos de dimensão quatro, aos quais se aplicam variáveis análogas ao caso tridimensional. Analisa-se também um exemplo de classificações de alunos (entre 1 e 5). Neste exemplo, existe também uma estrutura interna estabelecida pelas disciplinas e, por esse motivo, testa-se o efeito da ordenação das classificações.

Note-se que as classificações dos alunos (consideradas, neste caso, entre 1 e 5) podem ser vistas como dados categóricos ordenados.

Não obstante a escolha das variáveis ter sido, deliberadamente, preparada como um utilizador sem preocupações, parecem ter sido obtidos factos relevantes, como a classificação e organização dos indivíduos. Concluímos existirem três tipos de variáveis que se separam em blocos: as analíticas, as amplitudes (função dos extremos) e o máximo. A primeira componente é caracterizada pelas variáveis analíticas dado que estas são em maior número. As amplitudes definem a segunda componente principal e a terceira componente é caracterizada pela dimensão intermédia.

De acordo com o número de variáveis de cada grupo formam-se, na matriz de correlações, três blocos com mais ou menos variáveis.

Assim, por se organizarem em blocos, no capítulo VI, analisamos a organização espacial das variáveis e sua ligação com matrizes de correlações organizadas em blocos. Fizeram-se, experimentalmente, estudos para matrizes organizadas em dois e três blocos.

Como subproduto obtiveram-se resultados interessantes sobre os valores próprios de matrizes organizadas em blocos, sendo, contudo, necessário precisar os resultados (caso degenerado e estudo do break point, ou seja, quando uma variável se dissocia de um bloco para se associar a outro).

Como instrumento de auxílio à interpretação da ACP, ensaiou-se o uso de linhas de nível para as diferentes variáveis, o que se revelou satisfatoriamente eficaz. Estas linhas foram obtidas a partir do software *3DField*.

Também se analisou a estabilidade da ACP estudando a variação do primeiro valor para amostras de diferentes dimensões e a variação angular entre os três eixos principais do conjunto de todos os paralelepípedos e os eixos amostrais.

Todo o software para ACP foi construído de raiz em *MatLab*, pela maior facilidade de obtenção de um software aberto, que permitisse todo o tipo de modificações.

Simulações muito longas e sujeitas a pesado pós-tratamento foram levadas a cabo, pelo que se justifica a não realização de diversas tarefas que tiveram de ser adiadas para posteriores trabalhos (como por exemplo, um estudo mais alargado em relação à contextualização de dados categóricos).

Propositadamente, nesta fase este estudo foi experimental para tentar esclarecer o que é mais relevante.

Esta tese deve ser lida pensando no que os diferentes casos concretos dão e nas consequências que tem escolher mais ou menos variáveis de um tipo na formação da componente principal.

Capítulo II

Correlações com Outros Trabalhos

Em distintas áreas da ciência, ao serem levados a cabo estudos estatísticos consideram-se múltiplas variáveis para cada indivíduo. Resulta, muitas vezes, um extenso conjunto de dados que pode tornar-se penoso em termos de qualquer análise estatística. Assim, a redução da dimensão de um problema é um passo essencial antes de serem levados a cabo muitos tipos de análise.

Para esta redução, um critério geral consiste no interesse de preservar a estrutura fundamental dos dados originais e possibilitar a interpretação.

A análise em componentes principais constitui uma das ferramentas mais amplamente utilizadas na redução da dimensão de conjuntos de dados. Esta técnica obtém subespaços de pequena dimensão que tentam capturar o essencial da informação.

No entanto, alguns autores consideram que esta redução através de ACP não é real, uma vez que na formação das componentes principais estão envolvidas todas as variáveis originais. A este respeito McCabe ([4]) diz: ‘interpretação dos resultados e possíveis dados subsequentes e análise continuam a envolver todas as variáveis originais’. Por outro lado, é prática comum ignorar variáveis fracamente correlacionadas com as primeiras componentes principais. Cadima e Jolliffe mostraram que esta abordagem é potencialmente enganadora ([5],[6]).

Nesse sentido, estes autores desenvolveram uma estratégia de selecção de variáveis à priori de forma a reduzir a dimensão a partir das variáveis originais e não a partir das componentes principais [6]. Preocupam-se em distinguir um subconjunto das variáveis originais que é óptimo no sentido de aproximar as componentes principais relevantes, ou em alternativa, que é óptimo para uma aproximação dos dados originais. Em qualquer dos casos, a aproximação será considerada óptima em relação a determinados critérios pré-estabelecidos que dependem, obviamente, da natureza e do objectivo do estudo. Diversos critérios foram sugeridos por distintos autores, alguns deles referenciados em Duarte Silva [7] e Cadima, Minhoto e Cerdeira [8].

Contudo a análise de todos subconjuntos possíveis envolve um enorme esforço computacional. Note-se que para p variáveis originais podem ser definidos $2^p - 1$ subconjuntos distintos ([5],[6], [7],[8]).

Duarte Silva ([7]) desenvolveu uma metodologia computacional que permite analisar todos os subconjuntos de variáveis, considerando alguns dos mais importantes critérios de comparação propostos para dados. Um software para implementação pode ser encontrado em <http://www.porto.ucp.pt/psilva>. Este autor conclui que os algoritmos obtidos são praticáveis para problemas envolvendo até 30 variáveis.

Também Cadima, Minhoto e Cerdeira ([8]) estudaram aspectos computacionais de algoritmos construídos segundo diferentes critérios para optimização do subconjunto de variáveis a eliminar (*'to be discarded'*).

Métodos alternativos à ACP têm sido propostos. Autores como Vines, Rousson e Gasser sugerem a análise em componentes simples (SCA) ([9], [10]), cujo objectivo principal consiste numa melhoria da interpretação resultante de uma ACP. As componentes simples são uma aproximação das componentes principais, e caracterizam-se por poderem sempre ser representadas por inteiros. Esta técnica tem-se relevado vantajosa em termos práticos pela maior simplicidade. No entanto, representa uma menor quantidade de variabilidade extraída dos dados, para além de que as componentes simples estão muitas vezes correlacionadas

Existem, porém, aproximações diferentes ao problema da selecção de variáveis/redução da dimensão.

Gomes [11] e Figueiredo [12] seguiram a ideia de criação de um contexto para amostragem de variáveis. Procuram fazer a selecção *a priori* das variáveis, averiguando se é ou não redundante introduzir uma nova variável no estudo antes de fazer qualquer observação.

Enquanto na abordagem clássica, para um conjunto fixo de p variáveis, os dados representam uma amostra de n indivíduos, estes autores supõem que para um conjunto fixo de n indivíduos as p variáveis são seleccionadas, aleatoriamente, de uma determinada população de variáveis.

Pressupõem que as p variáveis seleccionadas aleatoriamente são centradas e reduzidas sobre a n -esfera e que se seguem uma distribuição de Bingham (Mardia, [13]) cujos parâmetros (direcção e concentração), desconhecidos em geral, podem ser estimados a partir de uma amostra de variáveis.

Em [12] são propostos testes de discordância que detectam a existência de um ou mais outliers numa amostra de variáveis proveniente de uma distribuição de Bingham sobre a n -esfera, permitindo a eliminação de uma ou mais variáveis. São também apresentadas regras de afectação de uma nova variável a grupos de variáveis previamente definidos e aos quais está associada um lei de Bingham.

Seguramente que todas estas metodologias são úteis, mas o problema analisado neste estudo é anterior a esse. Independentemente da contextualização, todos os autores referidos consideram as variáveis conhecidas de antemão, isto é, à partida a matriz de dados apresenta-se construída. E uma vez que a estrutura está definida também estão fixadas as componentes.

Neste trabalho o objectivo é diferente. Prende-se com a construção da matriz de dados e decisão sobre que variáveis escolher para ser utilizadas numa análise estatística (em particular, em ACP). Perante um conjunto de dados observados pretende-se saber se é ou não vantajoso em termos de interpretação construir novas variáveis a partir dos dados iniciais. Experimentam-se, ao acaso, diferentes combinações das variáveis básicas e analisa-se de que forma se organizam e contribuem para a formação dos eixos.

Capítulo III

Escolha de Variáveis num Contexto Geométrico

Neste capítulo vamos realizar análise em componentes principais sobre variáveis associadas a paralelepípedos rectângulos. Notar que este exemplo já foi tratado noutros contextos, como, por exemplo, em FC-FS [1].

3.1 O Exemplo de J. Davis, W.W. Cooley e P.R. Lohnes

Em FC-FS [1], seguindo J. Davis [2] e Cooley & Lohnes [3], são considerados 25 paralelepípedos $a \times b \times c$ gerados aleatoriamente, onde a , b e c seguem uma lei uniforme $[0,10]$ (anexo III).

A estes paralelepípedos aplicaram-se variáveis construídas a partir das arestas e escolhidas *‘de forma a permitir o estudo do ponto de vista psicológico das representações mentais de tamanho e de forma’*. Analisar-se-ão estas variáveis no capítulo V.

Os paralelepípedos gerados são divididos em três grandes grupos: *‘palitos’*, *‘tábuas’* e *‘cubos’* (Fig.3.1):



Fig. 3.1 – Tábuas (mica branca), cubos (‘bomba’ vulcânica de Timor) e Palitos (minério de cobre)

São designados por *‘palitos’* os paralelepípedos caracterizados por uma dimensão muito grande e as restantes muito pequenas. Por *‘tábuas’* denominam-se os paralelepípedos com duas dimensões muito grandes e uma dimensão muito pequena. Os *‘cubos’* são, naturalmente, os paralelepípedos com todas as dimensões iguais (ou quase).

Uma das ideias subjacentes a [1] é que qualquer indivíduo pode ser envolvido num paralelepípedo, de forma unívoca ou não.

Com efeito, no caso das classificações dos alunos, qualquer aluno pode ser visto como um hiper-paralelepípedo, onde cada aresta tem valor igual à classificação em cada disciplina. Ou melhor, o paralelepípedo pode ter a sequência de arestas seguindo a ordem das disciplinas e, neste caso, está definido de forma unívoca, ou no caso em que as classificações são ordenadas, a sequência de arestas ser arbitrária.

Também no caso das Águas de Gaia se podem considerar os consumidores como paralelepípedos multidimensionais onde cada aresta tem valor correspondente ao consumo em determinado bi-mês conhecido (definição unívoca). Se se ordenarem os valores dos consumos, a definição deixa de ser unívoca e os valores das arestas são arbitrários entre todos os consumos.

Embora o caso dos consumidores não seja uma situação tridimensional, pode-se igualmente estabelecer um paralelismo com os grupos de paralelepípedos considerados em [1]. Consumidores com consumos idênticos durante todo o ano serão hiper-cubos, consumidores com consumos elevados durante todo ano e um consumo pequeno (por exemplo, devido a férias) serão hiper-tábuas e consumidores com consumos pequenos durante todo o ano e um grande consumo serão hiper-palitos.

Para os alunos, evidentemente teremos hiper-cubos quando todas as notas foram iguais (ou muito semelhantes), hiper-palitos quando uma das notas for muito alta e as restantes baixas (o que acontece, por exemplo, com a disciplina de Educação Física: existem alunos que apenas apresentam boa classificação nesta disciplina) e hiper-tábuas quando todas as classificações são elevadas com exceção de uma disciplina (de novo, o exemplo de Educação Física: alunos com boas notas na maioria das disciplinas, não obtêm boa classificação nesta disciplina).

3.2 Paralelepípedos Rectângulos

Fixamos as dimensões dos paralelepípedos $a \times b \times c$ como inteiras, a variar entre 1 e 10, inclusive. Assim, vamos considerar 220 paralelepípedos distintos: $K_3^{10} = C_3^{10+3-1} = C_3^{12}$. Para cada paralelepípedo ordenaram-se as respectivas dimensões (por ordem crescente).

Considerou-se um conjunto de 12 variáveis:

X_1	<i>Comprimento da Aresta</i>	→	a
X_2	<i>Comprimento da Aresta</i>	→	b
X_3	<i>Comprimento da Aresta</i>	→	c
X_4	<i>Volume</i>	→	$X_1 X_2 X_3$
X_5	<i>Área Total</i>	→	$2X_1 X_2 + 2X_1 X_3 + 2X_2 X_3$
X_6	<i>Dimensão Máxima</i>	→	$Max(X_1, X_2, X_3)$
X_7	<i>Dimensão Mínima</i>	→	$Min(X_1, X_2, X_3)$
X_8	<i>Média das dimensões</i>	→	$Media(X_1, X_2, X_3)$
X_9	<i>Média das áreas laterais</i>	→	$Media(X_1 X_2, X_2 X_3, X_1 X_3)$
X_{10}	$\frac{\text{Máximo}}{\text{Mínimo}}$	→	$\frac{X_6}{X_7}$
X_{11}	<i>Máximo – Mínimo</i>	→	$X_6 - X_7$
X_{12}	$\frac{\text{Mínimo áreas laterais}}{\text{Máximo áreas laterais}}$	→	$\frac{Min(X_1 X_2, X_2 X_3, X_1 X_3)}{Max(X_1 X_2, X_2 X_3, X_1 X_3)}$

Tabela 3.1

Estas variáveis, dado serem geométricas, são facilmente interpretáveis.

As três primeiras são, naturalmente, as dimensões das arestas do paralelepípedo. Uma vez que se consideram as dimensões ordenadas, X_1 corresponde à dimensão mínima, X_2 à dimensão intermédia e a dimensão máxima é representada pela variável X_3 .

A quarta e quinta variável representam, respectivamente, o volume e área total. X_6 e X_7 são extremos. A média das dimensões corresponde a X_8 e X_9 à média das áreas laterais. X_{10} e X_{11} são medidas de cubicidade. X_{12} é uma combinação entre extremos de áreas laterais.

Portanto, consideram-se variáveis lineares, de superfície, de volume e, por outro lado, médias e extremos.

Inadvertidamente, dado a escolha de variáveis ser experimental, pode não se reparar que $X_9 = \frac{X_5}{6}$:

$$X_9 = \frac{X_1X_2 + X_1X_3 + X_2X_3}{3} = \frac{\frac{X_5}{2}}{3} = \frac{X_5}{6}.$$

Por outro lado, também $X_{12} = \frac{1}{X_{10}}$:

$$X_{12} = \frac{\min(X_1X_2, X_2X_3, X_1X_3)}{\max(X_1X_2, X_2X_3, X_1X_3)} = \frac{X_1X_2}{X_2X_3} = \frac{X_1}{X_3} = \frac{X_7}{X_6} = \frac{1}{X_{10}}.$$

Note-se também que, dado as dimensões estarem ordenadas dentro de cada paralelepípedo, $X_1 = X_7$ e $X_3 = X_6$.

3.3 Estatística Descritiva das Variáveis

Na tabela 3.2 encontram-se resumidas as estatísticas descritivas de cada uma das variáveis:

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Média	3,25	5,50	7,75	178,75	186,00	7,75	3,25	5,50	31,00	3,54	4,50	0,43
Mediana	3	5,5	8	108	161	8	3	5,5	26,834	2,5	4,5	0,4
Moda	1	5	10	36	88	10	1	5	14,667	2	4	0,5
Variância	4,41	5,88	4,41	36611	15248	4,41	4,41	3,26	423,56	6,47	5,88	0,06
Desvio-padrão *	2,10	2,42	2,10	191,34	123,48	2,10	2,10	1,81	20,58	2,54	2,42	0,25
Assimetria	0,88	0,00	-0,88	1,68	0,86	-0,88	0,88	0,00	0,86	1,24	0,00	0,57
Curtose **	3,0627	2,094	3,0627	5,7752	3,3131	3,0627	3,0627	2,559	3,3131	3,4636	2,094	2,3729
Mínimo	1	1	1	1	6	1	1	1	1	1	0	0,1
Máximo	10	10	10	1000	600	10	10	10	100	10	9	1
Coefficiente de Variação	0,65	0,44	0,27	1,07	0,66	0,27	0,65	0,33	0,66	0,72	0,54	0,58

* - Desvio Padrão Corrigido

** - Definição usada: $k = \frac{E(x-\mu)^4}{\sigma^4}$

Tabela 3.2

Na maior parte das variáveis não existem valores elevados de dispersão. O maior valor é apresentado pela variável X_4 (*Volume*), uma vez que se trata de um produto de grau três.

Tinha-se visto que $X_9 = \frac{X_5}{6}$. Logo, tem-se $\overline{X_9} = \frac{\overline{X_5}}{6} = \frac{186}{6} = 31$ e $s^2_{X_9} = \frac{s^2_{X_5}}{36} = \frac{15248}{36} = 423,556$, o que, efectivamente se verifica.

Para $X_{12} = \frac{1}{X_{10}}$ nada se pode concluir pois a relação existente não é linear.

Efectuam-se de seguida os ‘cálculos contínuos’ para a média e variância das variáveis X_1 e X_3 . Para tal supõe-se $X_1, X_2, X_3 \sim U_{[1,10]}$.

Tem-se que $X_3 = \text{Max}(X_1, X_2, X_3)$ e $X_1 = \text{Min}(X_1, X_2, X_3)$ e, dado que as distribuições das estatísticas ordinais são conhecidas, sabe-se que

$$F_{X_3}(x) = [F(x)]^n$$

$$F_{X_1}(x) = 1 - [1 - F(x)]^n$$

$$\text{Portanto, } F_{X_1}(x) = 1 - \left[1 - \left(\frac{x-1}{9}\right)\right]^3 = 1 - \left(\frac{10-x}{9}\right)^3 \text{ e } F_{X_3}(x) = \left(\frac{x-1}{9}\right)^3$$

Então tem-se que

$$E(X_1) = \int_{-\infty}^{+\infty} xf_{X_1}(x)dx = \int_1^{10} x \frac{1}{3} \left(\frac{10-x}{9} \right)^2 dx = 3.25$$

$$V(X_1) = E(X_1^2) - E^2(X_1) = \int_{-\infty}^{+\infty} x^2 f_{X_1}(x)dx - 3.25^2 = \int_1^{10} x^2 \frac{1}{3} \left(\frac{10-x}{9} \right)^2 dx = 3,04$$

$$E(X_3) = \int_{-\infty}^{+\infty} xf_{X_3}(x)dx = \int_1^{10} x \frac{3}{9^3} (x-1)^2 dx = 7.75$$

$$V(X_3) = E(X_3^2) - E^2(X_3) = \int_{-\infty}^{+\infty} x^2 f_{X_3}(x)dx - 7,75^2 = \int_1^{10} x^2 \frac{3}{9^3} (x-1)^2 dx = 3,04$$

etc.

Com estes resultados controlaram-se os valores da Tabela 3.2.

3.4 Análise em Componentes Principais

3.4.1 Matriz de Correlações

Segue-se a matriz de correlações (Tabela 3.3) para o conjunto de dados dos paralelepípedos, podendo verificar-se que entre as variáveis existem fortes correlações:

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1	0,58	0,33	0,91	0,86	0,33	1	0,77	0,86	-0,71	-0,58	0,83
X2	0,58	1	0,58	0,74	0,86	0,58	0,58	0,89	0,86	-0,21	0,00	0,28
X3	0,33	0,58	1	0,50	0,63	1,00	0,33	0,77	0,63	0,21	0,58	-0,18
X4	0,91	0,74	0,50	1	0,97	0,50	0,91	0,88	0,97	-0,49	-0,35	0,61
X5	0,86	0,86	0,63	0,97	1	0,63	0,86	0,97	1	-0,42	-0,20	0,52
X6	0,33	0,58	1	0,50	0,63	1	0,33	0,77	0,63	0,21	0,58	-0,18
X7	1	0,58	0,33	0,91	0,86	0,33	1	0,77	0,86	-0,71	-0,58	0,83
X8	0,77	0,89	0,77	0,88	0,97	0,77	0,77	1	0,97	-0,29	0,00	0,37
X9	0,86	0,86	0,63	0,97	1	0,63	0,86	0,97	1	-0,42	-0,20	0,52
X10	-0,71	-0,21	0,21	-0,49	-0,42	0,21	-0,71	-0,29	-0,42	1	0,79	-0,82
X11	-0,58	0,00	0,58	-0,35	-0,20	0,58	-0,58	0,00	-0,20	0,79	1	-0,87
X12	0,83	0,28	-0,18	0,61	0,52	-0,18	0,83	0,37	0,52	-0,82	-0,87	1

Tabela 3.3

Organiza-se a matriz de forma a obter blocos onde se concentrem as maiores correlações. Em particular, procura-se que as correlações existentes em cada bloco sejam sempre superiores, em módulo, a 50% :

	X5	X9	X8	X4	X1	X7	X2	X12	X10	X11	X3	X6
X5	1	1	0,97	0,97	0,86	0,86	0,86	0,52	-0,42	-0,2	0,63	0,63
X9	1	1	0,97	0,97	0,86	0,86	0,86	0,52	-0,42	-0,2	0,63	0,63
X8	0,97	0,97	1	0,88	0,77	0,77	0,89	0,37	-0,29	0	0,77	0,77
X4	0,97	0,97	0,88	1	0,91	0,91	0,74	0,61	-0,49	-0,35	0,5	0,5
X1	0,86	0,86	0,77	0,91	1	1	0,58	0,83	-0,71	-0,58	0,33	0,33
X7	0,86	0,86	0,77	0,91	1	1	0,58	0,83	-0,71	-0,58	0,33	0,33
X2	0,86	0,86	0,89	0,74	0,58	0,58	1	0,28	-0,21	0	0,58	0,58
X12	0,52	0,52	0,37	0,61	0,83	0,83	0,28	1	-0,82	-0,87	-0,18	-0,18
X10	-0,42	-0,42	-0,29	-0,49	-0,71	-0,71	-0,21	-0,82	1	0,79	0,21	0,21
X11	-0,2	-0,2	0	-0,35	-0,58	-0,58	0	-0,87	0,79	1	0,58	0,58
X3	0,63	0,63	0,77	0,5	0,33	0,33	0,58	-0,18	0,21	0,58	1	1
X6	0,63	0,63	0,77	0,5	0,33	0,33	0,58	-0,18	0,21	0,58	1	1

Tabela 3.4

Aparentemente, parecem formar-se três blocos, embora não sejam totalmente estanques uma vez que existem algumas correlações significativas interblocos. Note-se que como se pretendem blocos onde as correlações sejam sempre superiores a 50%, pode acontecer a exclusão de variáveis que embora possuindo correlações fortes com algumas variáveis de um bloco, não possuem com todas elas.

O bloco de maior dimensão é formado pelas variáveis X_5 (Área Total), X_9 (Média das áreas laterais), X_8 (Média das dimensões), X_4 (Volume), X_7 (Dimensão Mínima) e pelas dimensões X_1 e X_2 , ou seja, é constituído por variáveis analíticas, mínimo e dimensão intermédia.

Em relação às correlações as variáveis deste bloco e as restantes destaca-se a sua correlação com $X_3 (= X_6)$, com excepção de $X_1 (= X_7)$. Também com X_{12} apresentam algumas correlações significativas.

Por outro lado, X_{10} , X_{11} e X_{12} formam o segundo maior bloco: são funções dos extremos, em particular, representam amplitudes. X_{12} relaciona-se negativamente com X_{10} e X_{11} e estas duas relacionam-se positivamente. O sentido da correlação é natural dado que $X_{12} = \frac{1}{X_{10}}$. Com as restantes variáveis, a destacar apenas a correlação com $X_1 (= X_7)$.

A variável Máximo $X_6 (= X_3)$ encontra-se isolada num terceiro bloco.

3.4.2 Valores Próprios da Matriz de Correlações

Na tabela seguinte estão apresentados os valores próprios obtidos para a matriz de correlações:

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
7,587	63,23	63
3,55	29,57	93
0,46	3,81	97
0,25	2,07	99
0,12	1,01	100
0,03	0,29	100
0,00	0,02	100
0,00	0,00	100
0,00	0,00	100
0,00	0,00	100
0,00	0,00	100
0,00	0,00	100

Tabela 3.5

Verifica-se que a primeira componente principal explica 59,73% da variância total. As duas primeiras componentes, em conjunto, explicam 90,46% da inércia total. Consequentemente, duas componentes principais podem ser consideradas suficientes para resumir a informação total.

3.4.3 Interpretação das Componentes Principais

No que se segue, designa-se por (r, s) o plano das componentes r e s : r horizontal e s horizontal.

Cada paralelepípedo de dimensões a, b e c é representado por $(a-1)(b-1)(c-1)$. Assim, por exemplo, 059 representa o paralelepípedo $1 \times 6 \times 10$.

Apresenta-se na Fig.3.4 a representação gráfica dos indivíduos no plano formado pelos dois primeiros eixos principais de inércia (plano (1,2)).

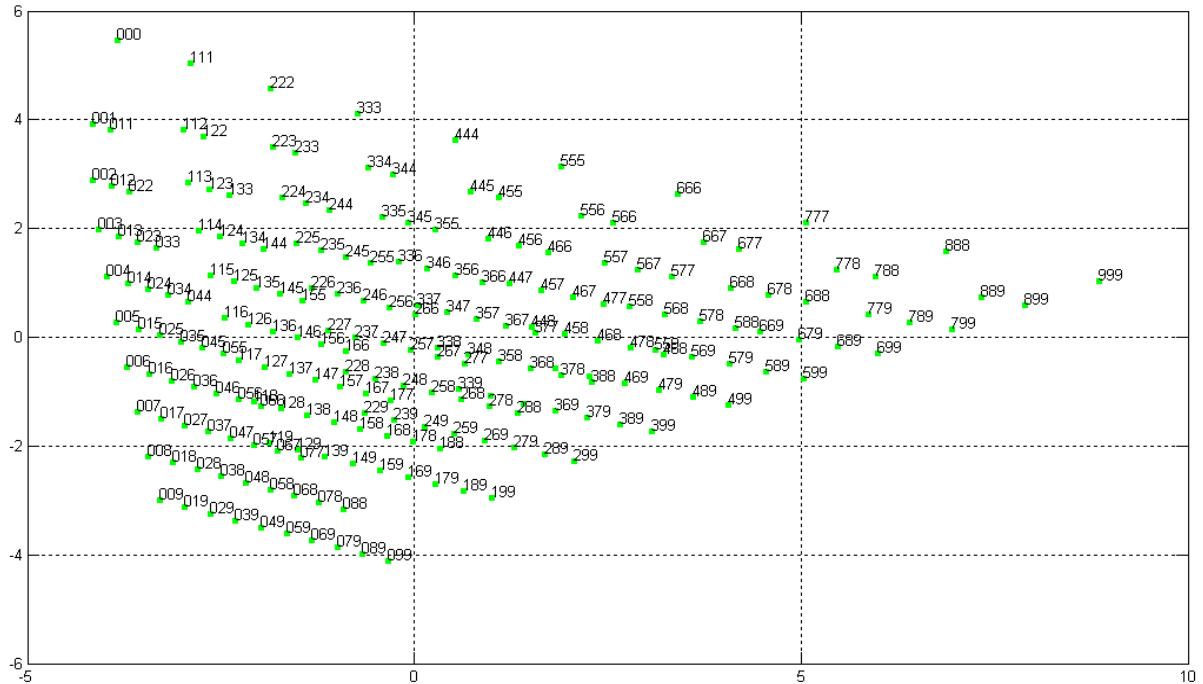


Fig.3.4 – Indivíduos representados no plano principal (1,2)

Observa-se que os dados se organizam segundo um trapézio, praticamente um triângulo isósceles com um dos ângulos menos agudo truncado. Este triângulo com outra orientação já existe em [1], mas não foi detectado explicitamente.

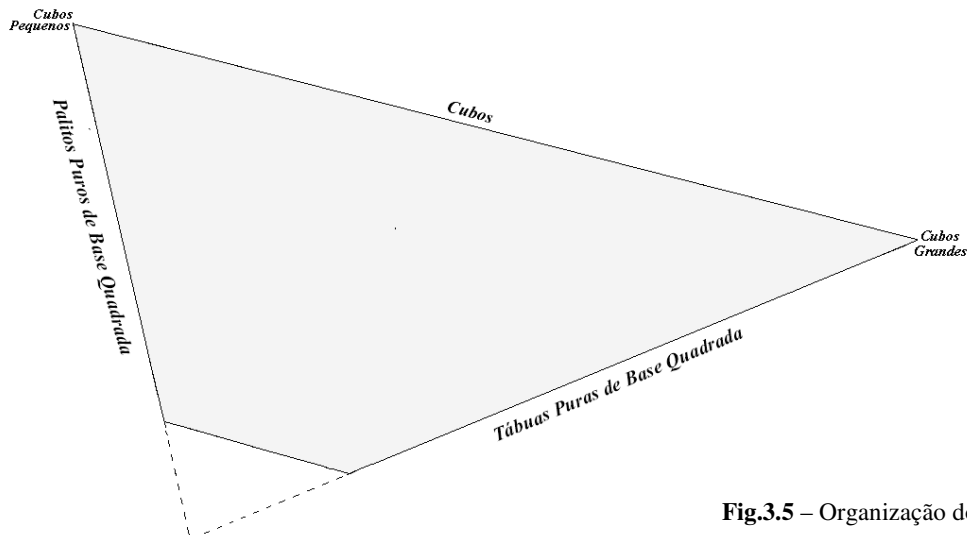


Fig.3.5 – Organização dos Indivíduos

Verifica-se que a primeira componente tem a direcção do vértice mais agudo e está fortemente ligada ao volume. De facto, para o volume ser grande têm que ser grandes as três arestas, isto é, o paralelepípedo tem de ser, aproximadamente, um cubo.

A segunda componente está ligada ao afastamento da cubicidade, em particular à relação entre o máximo e o mínimo, e em que a dimensão intermédia pouco intervém.

Constata-se que os lados do trapézio têm uma interpretação simples à excepção da base menor.

A base maior é constituída pelos cubos, sendo que à medida que nos deslocamos no sentido do vértice mais agudo os cubos vão tomando maiores dimensões.

Na aresta à esquerda, e tomando o sentido do vértice truncado do triângulo, cubos de pequenas dimensões aproximam-se de palitos puros de base quadrada. Na outra aresta, e considerando o sentido do vértice mais agudo, tábuas puras de base quadrada vão-se aproximando dos cubos de grandes dimensões.

A base menor é uma zona de transição entre palitos e tábuas.

Note-se que com apenas duas componentes a ACP organizou os paralelepípedos.

A interpretação dos eixos fica assim muito clara.

3.4.3.1 Interpretação da Primeira Componente Principal

No primeiro eixo observam-se os seguintes grupos em oposição:

Grandes Cubos	999	vs	000	Pequenos 'Quase-Cubos'	Cubos e 'Palitos'
	899		001		
	889		011		
	888		002		
	799		012		
	789		022		
	788		003		
	779		013		
	778		023		
	699		004		
	689		005		
	778		015		
			006		
			016		
			007		
			008		

Tabela 3.6

No primeiro grupo encontram-se cubos de grandes dimensões. Em oposição, no 2º grupo, encontram-se paralelepípedos do tipo '*palito*' e também '*quase cubos*' de pequenas dimensões. Todos os paralelepípedos deste grupo têm pelo menos duas dimensões pequenas.

Observando o círculo de correlações (Fig.3.6) verifica-se que todas as variáveis estão muito próximas do bordo, o que se compreende pela geometria das relações, mas o que também significa correlações muito fortes. Ou seja, muitas variáveis são redundantes e podem ser resumidas em poucas componentes.

As duas variáveis que se encontram mais no interior do círculo são claramente funções não lineares das arestas e estão ligadas a um aspecto menos marcado que é a variável intermédia: uma pela sua ausência (X_{10}), outra pela sua presença (X_2).

As variáveis que mais contribuem para a formação da primeira componente são, claramente, variáveis pertencentes, na matriz de correlações, ao bloco de maior dimensão.

	Y1
X1	0,94
X2	0,80
X3	0,56
X4	0,96
X5	0,98
X6	0,56
X7	0,94
X8	0,94
X9	0,98
X10	-0,55
X11	-0,33
X12	0,65

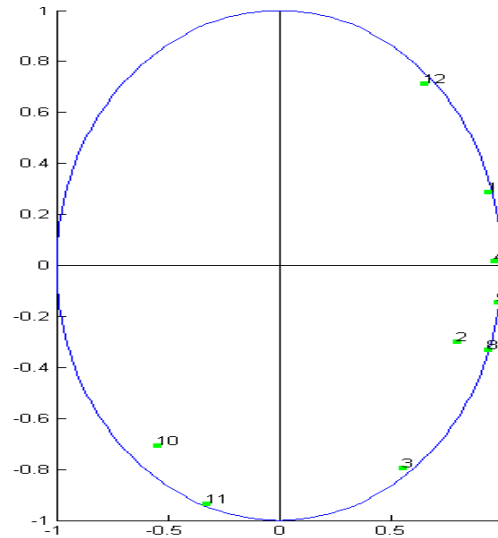


Fig.3.6 - Círculo de correlações (12)

Note-se que como $X_5 = 6X_9$, $X_3 = X_6$ e $X_1 = X_7$ não se repetiu a representação no círculo de correlações.

De entre todas as variáveis as que mais contribuem para a sua formação são as variáveis *volume*, *área total* e *dimensão mínima*:

Analisando o volume e área lateral dos paralelepípedos em grande oposição conclui-se que os paralelepípedos do primeiro grupo apresentam os maiores volumes e a maior área total, enquanto que aos paralelepípedos do segundo grupo representam os menores valores nestas variáveis (Tabela 3.7):

Paralelepípedo	Volume	Área Total		Paralelepípedo	Volume	Área Total
999	1000	600	vs	000	1	2
899	900	560		001	2	10
889	810	522		011	4	16
888	729	486		002	3	14
799	720	520		012	6	22
789	800	484		022	9	30
788	640	450		003	4	18
779	648	448		013	8	28
699	700	480		023	12	34
689	630	446		004	5	22
778	648	416		005	6	26
				015	12	40
				006	7	30
				016	14	46
				007	8	34
				008	9	38

Tabela 3.7

Estas conclusões são bastante visíveis traçando as linhas equi-volume (Fig.3.7) e equi-área (Fig.3.8). Assinale-se que estas linhas não estão igualmente espaçadas. O valor das linhas de nível do volume e da área vai aumentando no sentido positivo do 1º eixo, ou seja, do segundo grupo de paralelepípedos para o primeiro grupo:

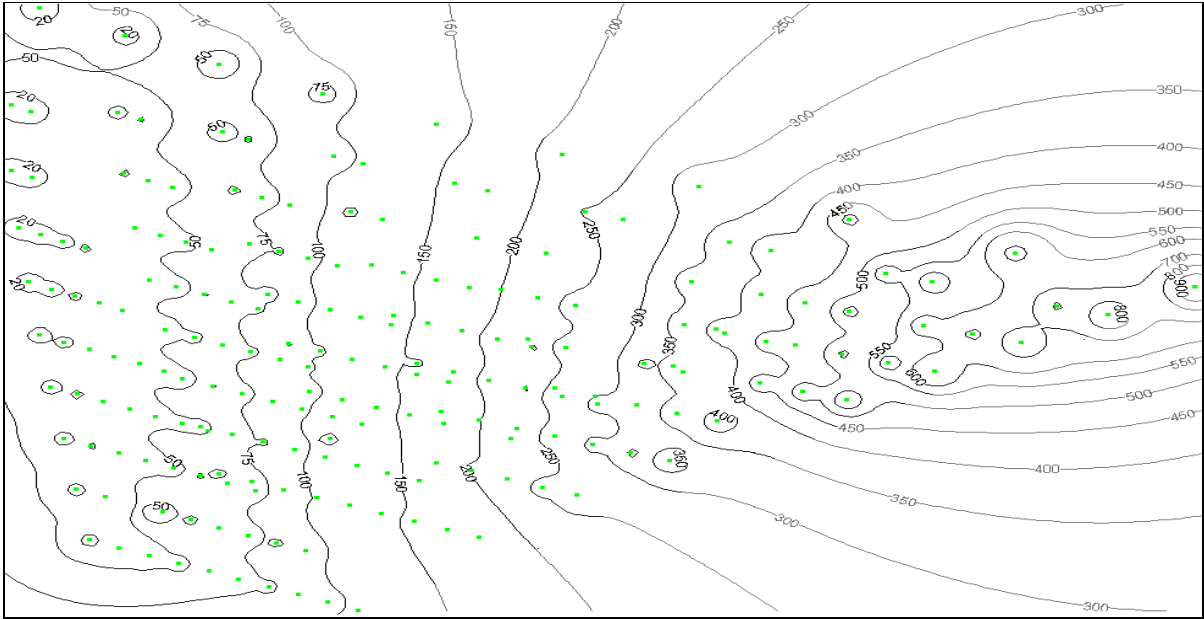


Fig.3.7 – Linhas equi-volume (12)

Repare-se que a zona dos pequenos volumes (à esquerda) é um ‘vale’ profundo, não um simples ‘poço’ enquanto a zona dos grandes volumes é um ‘pico escarpado’. As linhas de nível na parte esquerda são praticamente verticais, começando a envolver o pico na parte direita.

Nas linhas equi-área, abaixo apresentadas, verifica-se que não existe um pico tão ‘escarpado’, nem um ‘vale’ tão profundo. Ou seja, a área distingue menos os tamanhos.

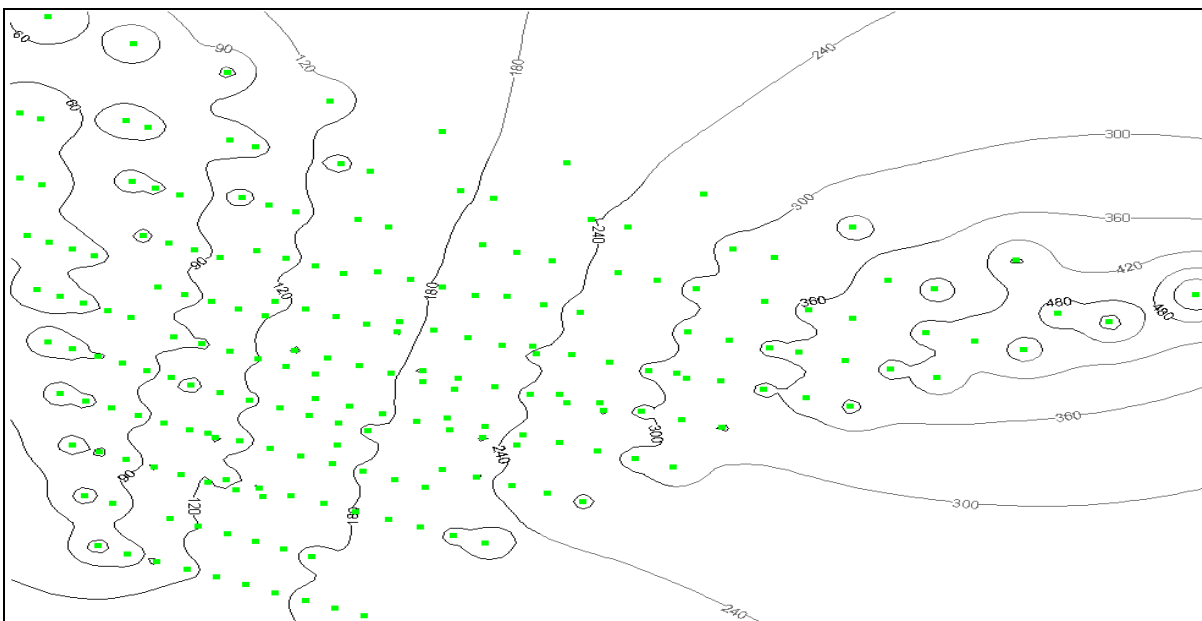


Fig.3.8 – Linhas equi-área (12)

As linhas de nível são verticais no domínio (trapézio).

Assim, a primeira componente é o tamanho (volume ou área total).

3.4.3.2 Interpretação da Segunda Componente Principal

No segundo eixo encontram-se em oposição os seguintes grupos de paralelepípedos:

000		099
111		089
222		059
333		079
001		039
011	vs	088
112		069
122		029
223		049
444		
332		

Tabela 3.8

Opõem-se paralelepípedos que tem pelo menos duas componentes iguais (e pequenas) com os que não têm, isto é, têm pelo menos duas muito diferentes (o que implica pelo menos uma grande e uma pequena). Neste segundo grupo, será a dimensão intermédia que determinará se se tratam de *‘palitos’* ou *‘tábuas’*.

Relativamente à segunda componente as variáveis parecem dividir-se em três grupos: variáveis com correlações negativas (correlações próximas de -1), uma com correlação positiva e as que possuem correlação quase nula. Ou seja, podem definir-se os seguintes conjuntos de variáveis: $\{X_3, X_6, X_{10}, X_{11}\}$, $\{X_{12}\}$ e $\{X_1, X_2, X_4, X_5, X_7, X_8, X_9\}$.

	Y2
X1	0,29
X2	-0,30
X3	-0,79
X4	0,02
X5	-0,14
X6	-0,79
X7	0,29
X8	-0,33
X9	-0,14
X10	-0,71
X11	-0,94
X12	0,71

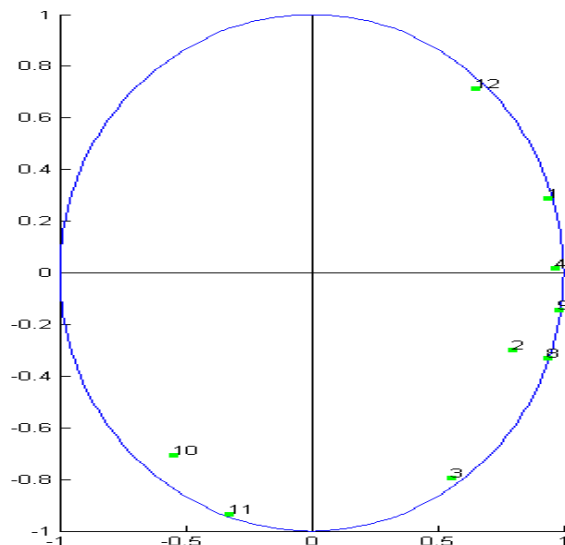


Fig.3.9 – Círculo de correlações (12)

A análise das correlações mostra que a variável *Máximo – Mínimo* é preponderante para a formação desta componente. Os paralelepípedos do segundo grupo apresentam os maiores valores nesta variável já que se caracterizam por ter uma dimensão pequena e outra grande.

Observando as linhas equi-Máximo-Mínimo (Fig.3.10) verifica-se que o valor que estas representam aumenta no sentido negativo do 2º eixo, ou seja, do primeiro grupo de paralelepípedos para o segundo, com as linhas de nível paralelas ao 1º eixo (horizontais).

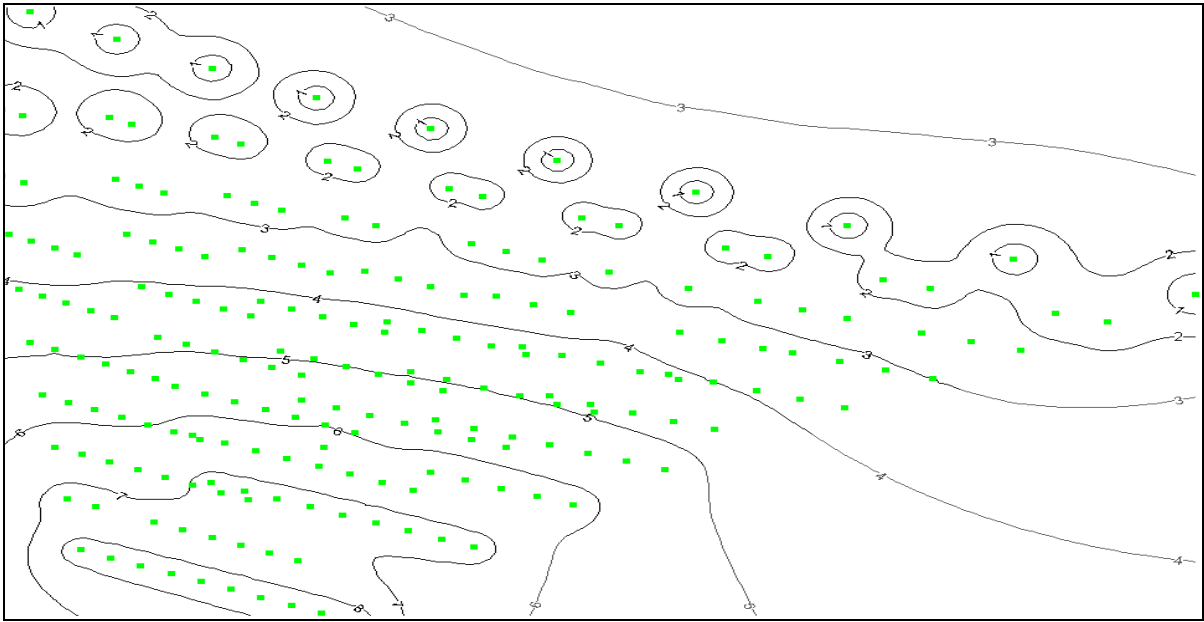


Fig. 3.10 - Linhas equi- Máximo-Mínimo (12)

Também a variável X_{10} possui um papel determinante na formação desta componente.

Tem-se que as variáveis mais fortemente correlacionadas com o 2º eixo pertencem, na matriz de correlações, aos 2º e 3º blocos. A variável X_{11} (*Máximo-Mínimo*), com maior importância para a segunda componente, pertence ao segundo bloco de maior dimensão.

3.4.3.3 Interpretação da Terceira Componente Principal

Segue-se a representação gráfica dos indivíduos no plano principal (1,3):

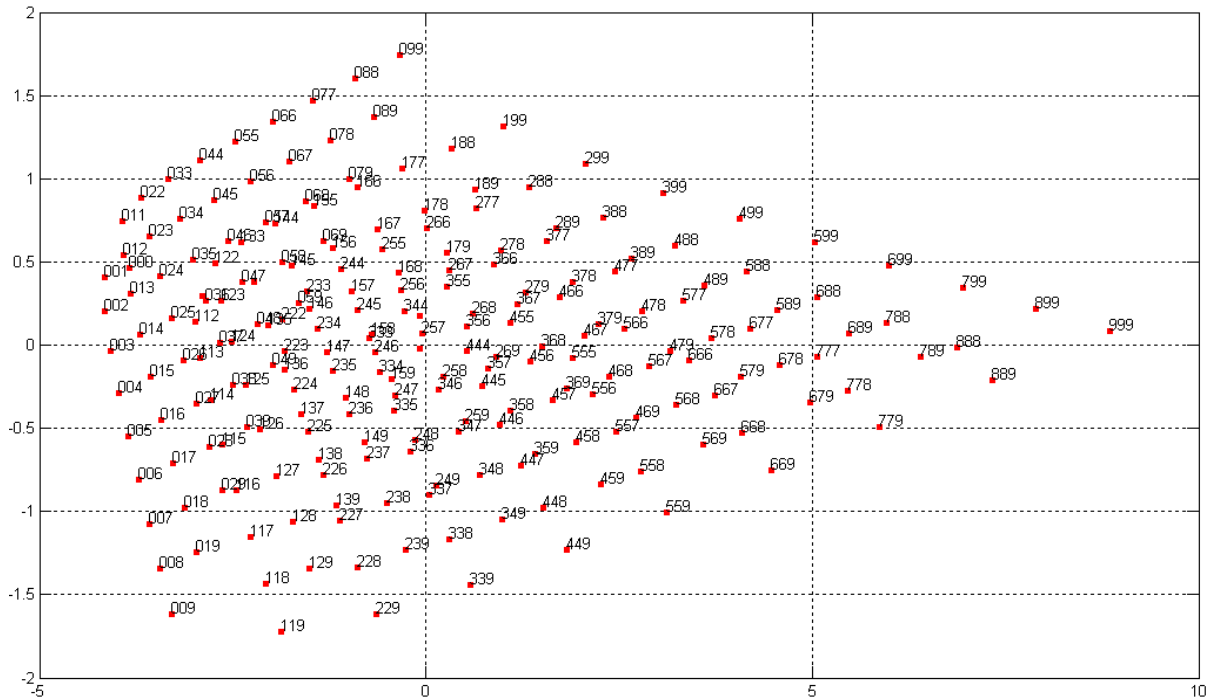


Fig.3.11 – Indivíduos representados no plano principal (1,3)

No terceiro eixo observam-se, como é visível na figura 3.12 e Tabela 3.9, os seguintes grupos em oposição:

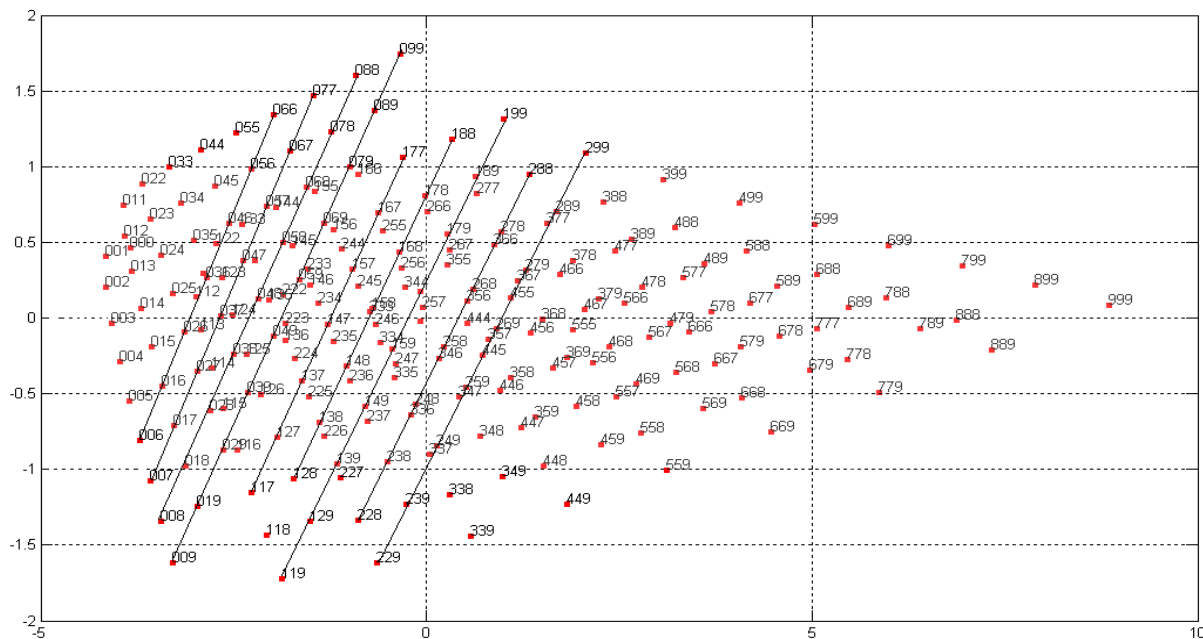


Fig.3.12 – Grupos em Oposição no terceiro eixo (plano (1,3))

009		099
119		088
229		077
118	vs	066
339		089
008		199
129		055
228		078

Tabela 3.9

Este eixo parece distinguir paralelepípedos do tipo ‘tábua’ dos paralelepípedos tipo ‘palito’.

Todos os paralelepípedos acima referidos possuem uma dimensão grande e outra pequena. No entanto, a dimensão intermédia (X_2) do primeiro grupo é pequena enquanto no segundo grupo é grande. Tem-se assim ‘palitos’ vs ‘tábuas’, respectivamente.

O círculo de correlações revela a existência de correlações fracas entre esta componente e as variáveis. Com exceção da variável X_2 , com correlação de 50%, e sendo portanto, a que mais contribui para a formação deste eixo, todas as outras variáveis possuem correlação inferior a 25% (Fig.3.13).

	Y3
X1	-0,17
X2	0,50
X3	-0,23
X4	0,02
X5	0,09
X6	-0,23
X7	-0,17
X8	0,07
X9	0,09
X10	0,11
X11	-0,05
X12	-0,07

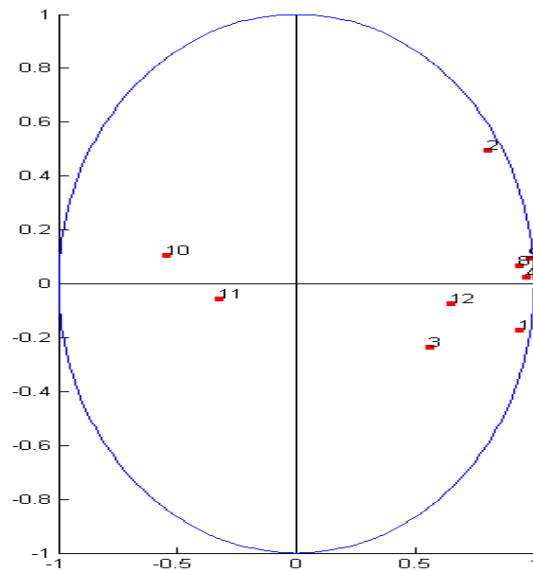


Fig.3.13 -Círculo de correlações (13)

A terceira componente, sendo muito importante por ter fácil interpretação (distingue entre palitos e tábuas) não é muito marcada dada a pouca inércia explicada. E as linhas de nível que seriam horizontais encontram-se pouco evidentes.

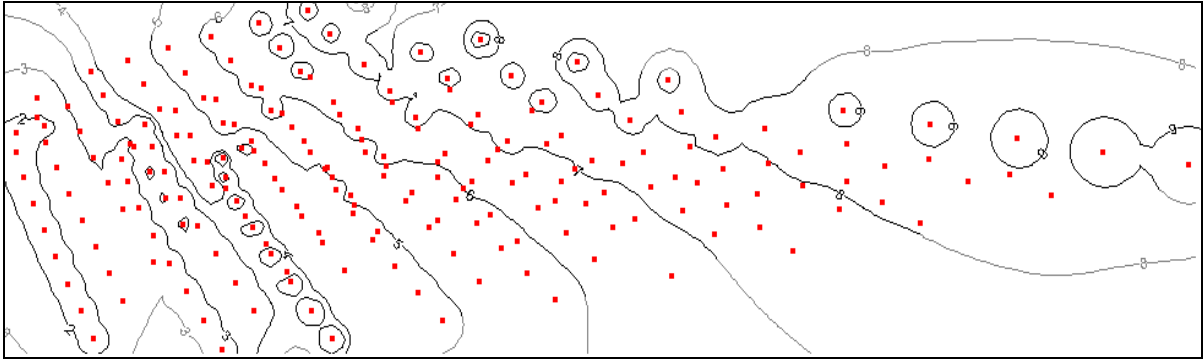


Fig.3.14 – Linhas equi-Dimensão Intermédia (13)

3.5 Estabilidade da ACP Realizada

3.5.1 Variação do primeiro valor próprio em função da dimensão

Analisa-se a influência da dimensão da amostra no primeiro valor próprio e, consequentemente, da quantidade da informação explicada pela primeira componente.

Recolheram-se, ao acaso, 30 amostras de dimensões sucessivamente 10, 20, 30, ... 210. Cada amostra foi obtida sem reposição.

Seguem-se os resultados obtidos para o primeiro valor próprio da matriz de correlações correspondente a cada amostra.

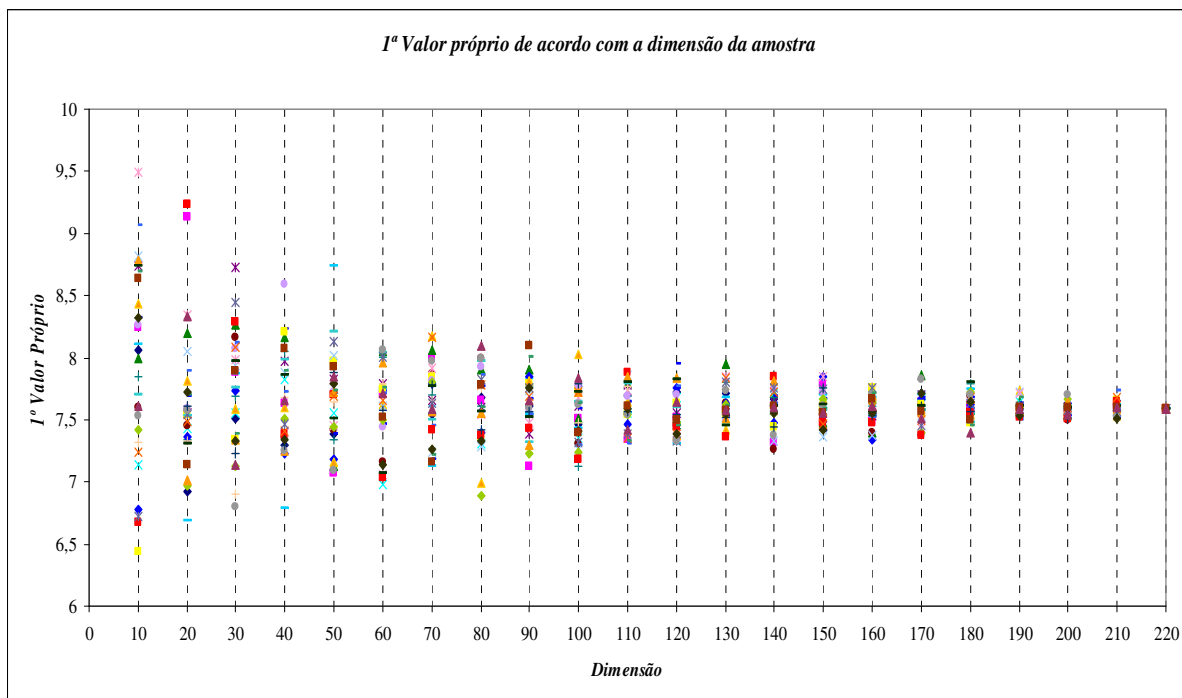


Fig.3.15 -Variação do primeiro valor próprio em função da dimensão

Os gráficos demonstram que a variação do primeiro valor próprio diminui com o aumento da dimensão.

Designa-se por valor assintótico o valor próprio correspondente aos 220 paralelepípedos ($\lambda = 7,589$).

Se, por qualquer razão, se tiver uma amostra aleatória verifica-se o seguinte: alguma tendência para o primeiro valor próprio ser excessivo nas amostras pequenas e muito rápida aproximação do valor assintótico com o aumentar da dimensão da amostra.

Através do diagrama caixa de bigodes (Fig.3.16) verifica-se que para amostras de pequena dimensão, em particular $n = 10, 20$ e 30 , as flutuações são ainda importantes. O coeficiente de variação oscila entre 7 e 10% (Tabela 3.10).

No entanto, para amostras maiores, as flutuações diminuem, sendo que o para amostras de tamanho maior que 100, o coeficiente de variação é sempre menor ou igual a 2%.

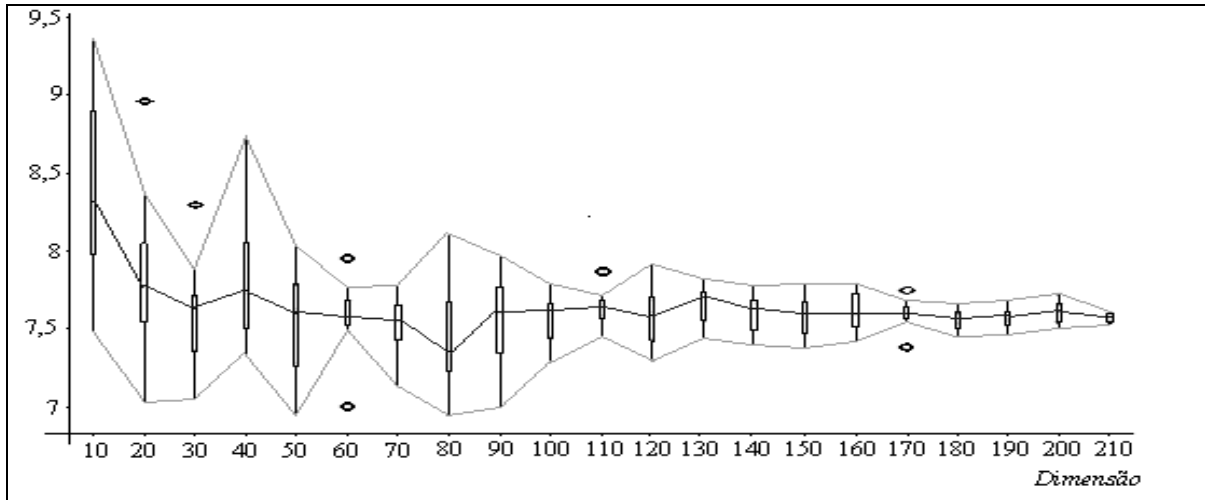


Fig.3.16 – Diagrama caixa de bigodes

Análise da normalidade do 1º valor próprio

Estudou-se, por simulação, o estimador do primeiro valor próprio:

Dimensão	Média	Desvio Padrão	Assimetria	Curtose	Coefficiente de Variação
10	7,9680	0,7726	-0,2324	2,2833	0,0970
20	7,6506	0,5686	1,1449	4,5562	0,0743
30	7,6612	0,5906	-1,1170	5,3518	0,0771
40	7,6671	0,3916	0,1918	2,7306	0,0511
50	7,6469	0,3705	0,6190	3,7803	0,0485
60	7,6111	0,3217	-0,5573	2,3785	0,0423
70	7,6482	0,2972	-0,1456	2,1786	0,0389
80	7,6132	0,2906	-0,5387	2,9570	0,0382
90	7,6098	0,2235	-0,0059	2,7800	0,0294
100	7,5389	0,2198	0,0215	2,2863	0,0292
110	7,6103	0,1696	-0,1888	2,0293	0,0223
120	7,5502	0,1701	0,5796	2,4622	0,0225
130	7,6160	0,1369	0,5269	2,7842	0,0180
140	7,5948	0,1496	-0,5328	2,6636	0,0197
150	7,5983	0,1268	0,2345	2,4153	0,0167
160	7,5825	0,1146	-0,3070	2,4239	0,0151
170	7,5991	0,1218	-0,0322	2,6638	0,0160
180	7,6064	0,0996	0,0330	2,3859	0,0131
190	7,5943	0,0601	0,8187	3,0812	0,0079
200	7,5741	0,0512	0,5526	2,6979	0,0068
210	7,5902	0,0482	0,7772	4,0575	0,0064

Tabela 3.10

A análise da normalidade do primeiro valor próprio foi feita aplicando o teste de normalidade (γ_1, γ_2) ou de Kendall & Stuart (Arvazian, [13]).

Segundo este teste, ao nível de significância de 5%, aceita-se o carácter normal caso se verifique

$$\begin{cases} |\hat{\gamma}_1(n)| < 1,5\sigma_{\hat{\gamma}_1} \\ \left| \hat{\gamma}_2(n) - 3 + \frac{6}{n+1} \right| < 1,5\sigma_{\hat{\gamma}_2} \end{cases}$$

e rejeita-se caso

$$\begin{cases} |\hat{\gamma}_1(n)| \geq 2\sigma_{\hat{\gamma}_1} \\ \left| \hat{\gamma}_2(n) - 3 + \frac{6}{n+1} \right| \geq 2\sigma_{\hat{\gamma}_2} \end{cases},$$

onde $\hat{\gamma}_1(n)$ e $\hat{\gamma}_2(n)$ representam a assimetria e curtose amostrais, respectivamente.

Neste caso, com $n = 30$, decide-se de acordo com a figura 3.17:

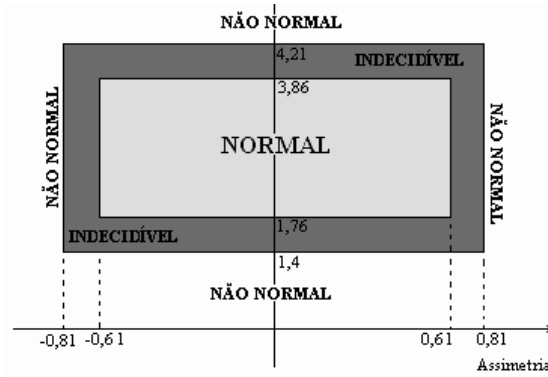


Fig.3.17

Observando a representação Assimetria vs Curtose (Fig.3.18) para as diferentes dimensões, verifica-se que a distribuição amostral do 1º valor próprio passa testes de normalidade, com excepção de $n = 20$, $n = 30$ e $n = 190$. Para $n = 50$ e $n = 210$ o teste não decide.

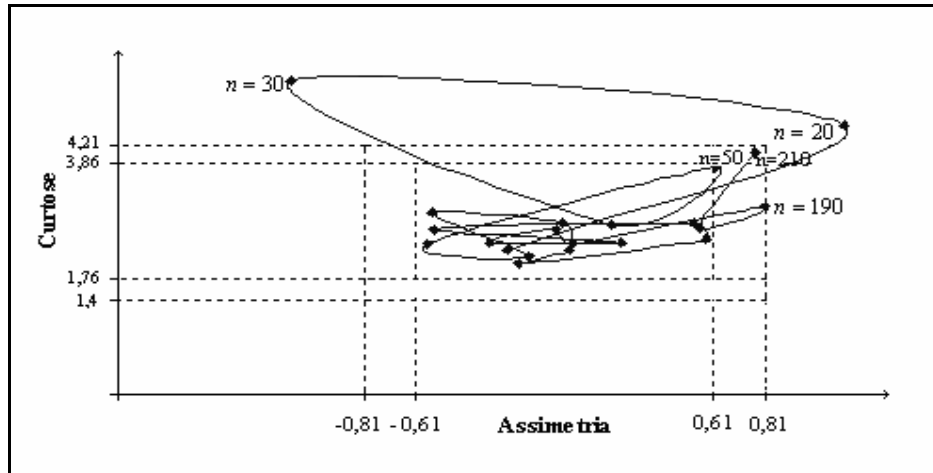


Fig.3.18 - Representação Assimetria vs Curtose

Os casos são extremos na amostragem, isto é, amostras que contêm quase todos os paralelepípedos ou amostras pequenas.

Note-se que na maior parte das amostras, mesmo com todas as não normalidades existentes, o primeiro valor próprio passa quase sempre o teste de normalidade (γ_1, γ_2) .

De facto, apenas 5 valores em 21 (25%) não passam o teste de normalidade.

Examina-se agora *a variação do primeiro valor em função da dimensão da amostra quando se retira uma variável ou grupos de variáveis*.

Verifica-se que o primeiro valor próprio sofre alterações para cada variável retirada, sendo tanto maior a alteração quanto maior a importância da variável em questão para a formação da primeira componente. Contudo, a sua variação para as diferentes dimensões é idêntica à ocorrida quando se tomam todas as variáveis.

A título de exemplo, apresenta-se o que ocorre retirando a variável X_4 , dado ser a variável com maior preponderância para a formação da primeira componente.

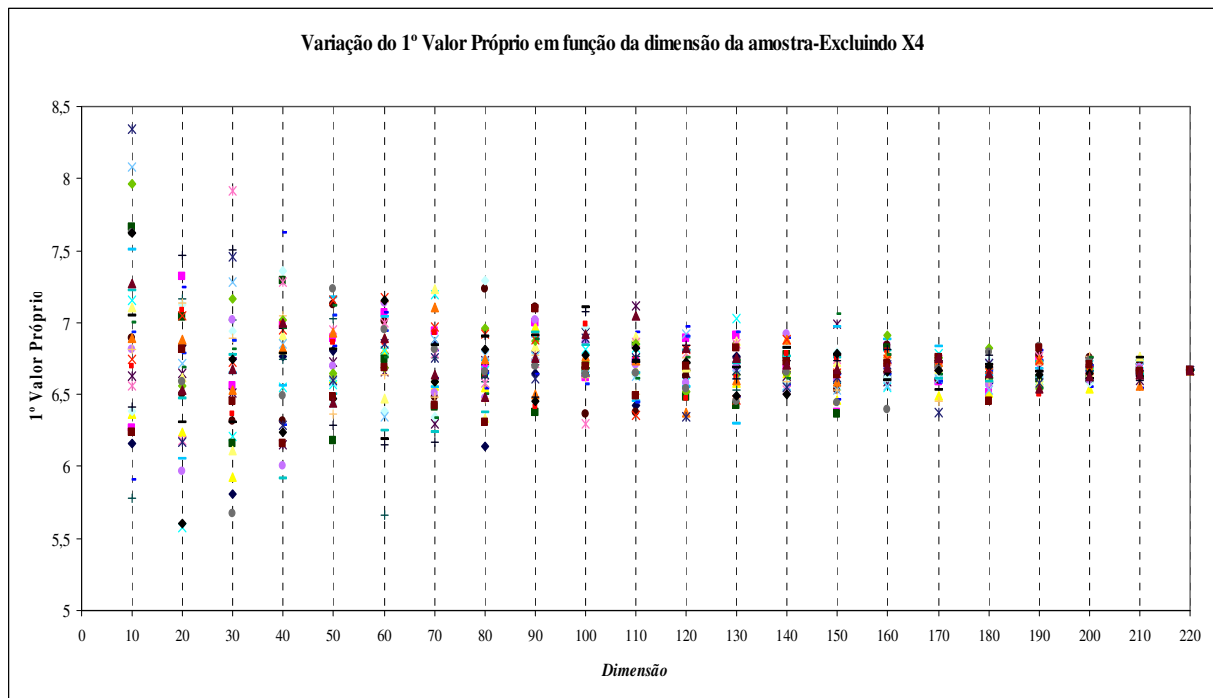


Fig.3.19-Variação do primeiro valor próprio em função da dimensão

As diferenças não parecem ser significativas, como se verifica analisando o diagrama caixa de bigodes de ambas as situações (Fig.3.20):

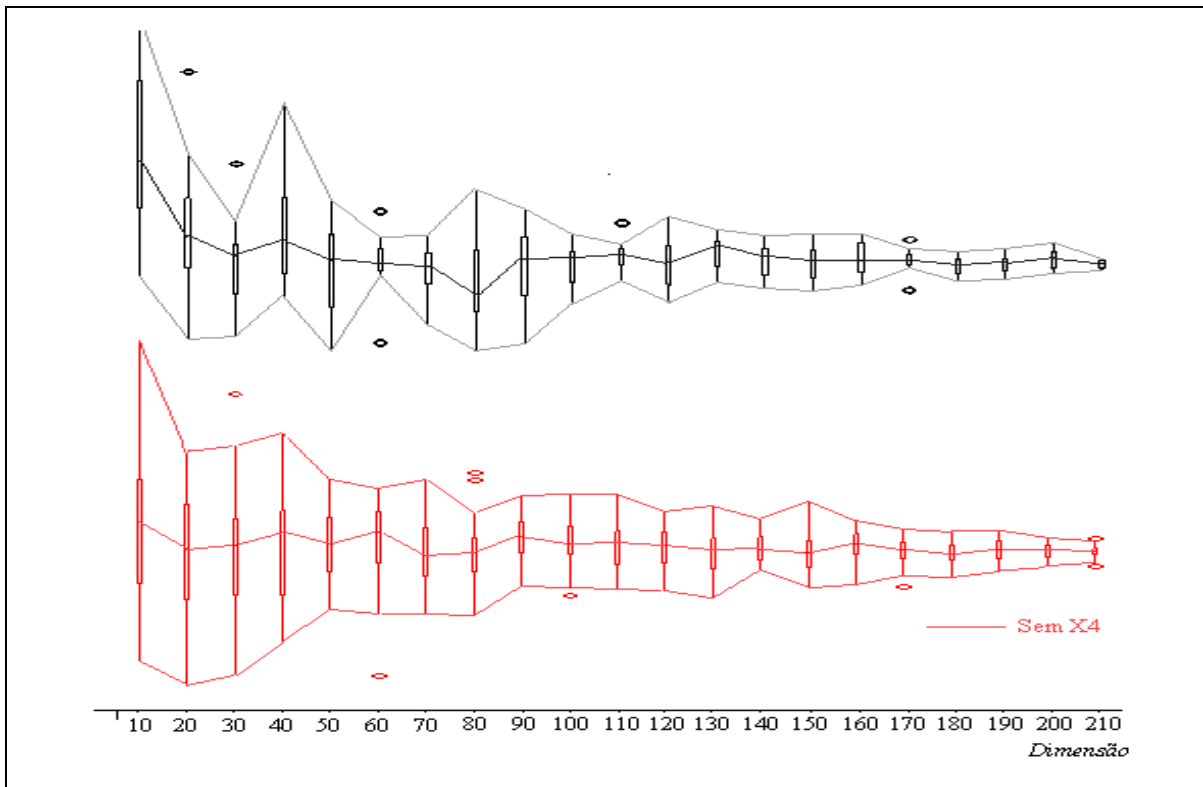


Fig.3.20 - Comparação de Diagramas Caixas de Bigodes

Aliás, se se compararem os valores médios do primeiro valor próprio para cada dimensão, quer com todas as variáveis, quer excluindo X_4 , observa-se o paralelismo na aproximação ao valor assintótico:

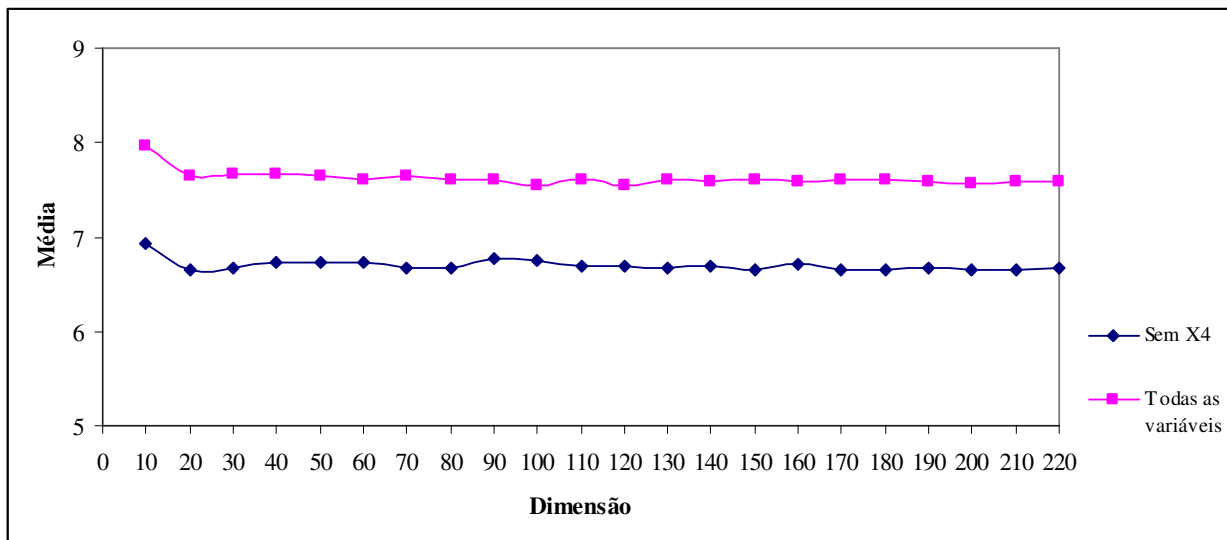


Fig.3.21 - Comparação de Valores Médios do Primeiro Valor Próprio

Quando se retiram *grupos de variáveis* verifica-se que a variabilidade do primeiro valor próprio é menor, embora também diminua com o aumento da amostra.

Segue a variação do primeiro valor próprio em função da dimensão quando retiradas as variáveis correspondentes ao bloco de maior dimensão na matriz de correlações:

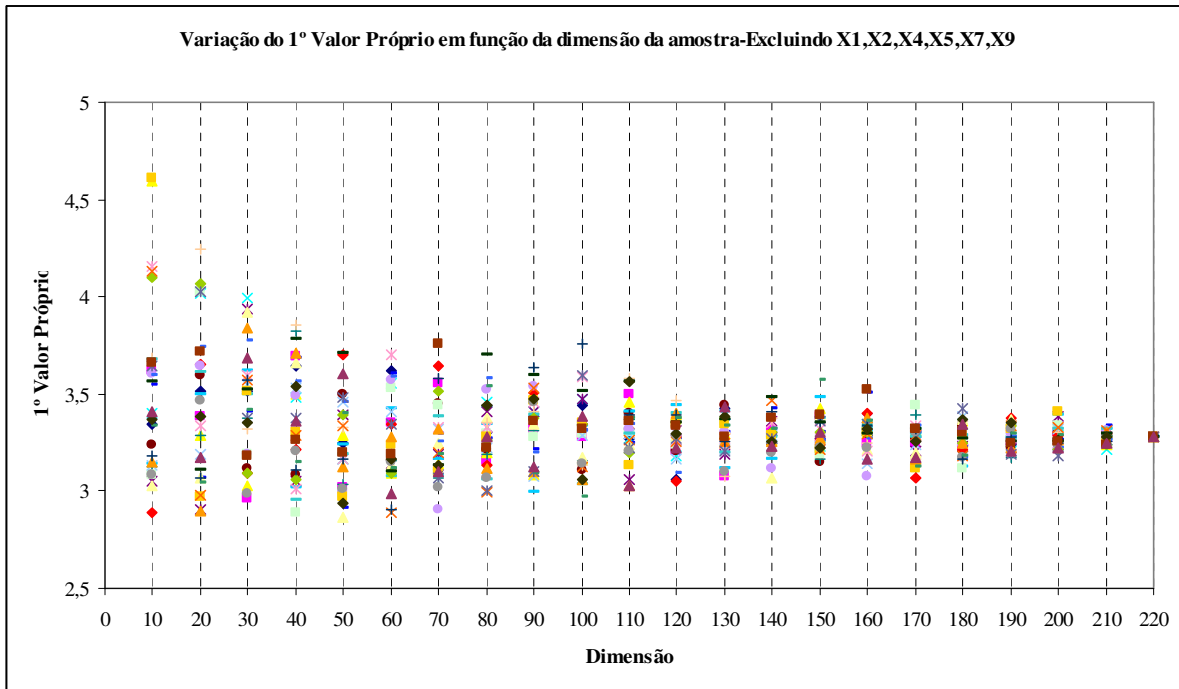


Fig.3.22-Variação do primeiro valor próprio em função da dimensão

Há, naturalmente, uma grande diminuição do valor assintótico ($\lambda = 3,2826$), mas na sua aproximação o comportamento é análogo ao que acontece para todas as variáveis.

Pode-se, portanto, concluir-se que a estrutura dos valores próprios é razoavelmente estável.

3.5.2 Variações Angulares

Como varia, em função da dimensão, o ângulo entre os diferentes eixos do conjunto de todos os paralelepípedos e o ‘eixos amostrais’?

A obtenção do ângulo é feita tendo em conta que o n-ésimo vector próprio é o vector director do n-ésimo eixo.

Note-se que $\cos(\hat{u,v}) = \frac{|\vec{u} \cdot \vec{v}|}{\|\vec{u}\| \|\vec{v}\|}$, onde cada vector representa o vector director do eixo real e do eixo

amostral.

Tal como no estudo da variação do primeiro valor próprio, recolheram-se, ao acaso, 30 amostras de dimensões sucessivamente 10, 20, 30, ... 210.

3.5.2.1 Primeira Componente

O gráfico seguinte refere-se à variação do ângulo entre o primeiro eixo real e o ‘primeiro eixo amostral’ (1º Ângulo) em função da dimensão da amostra.

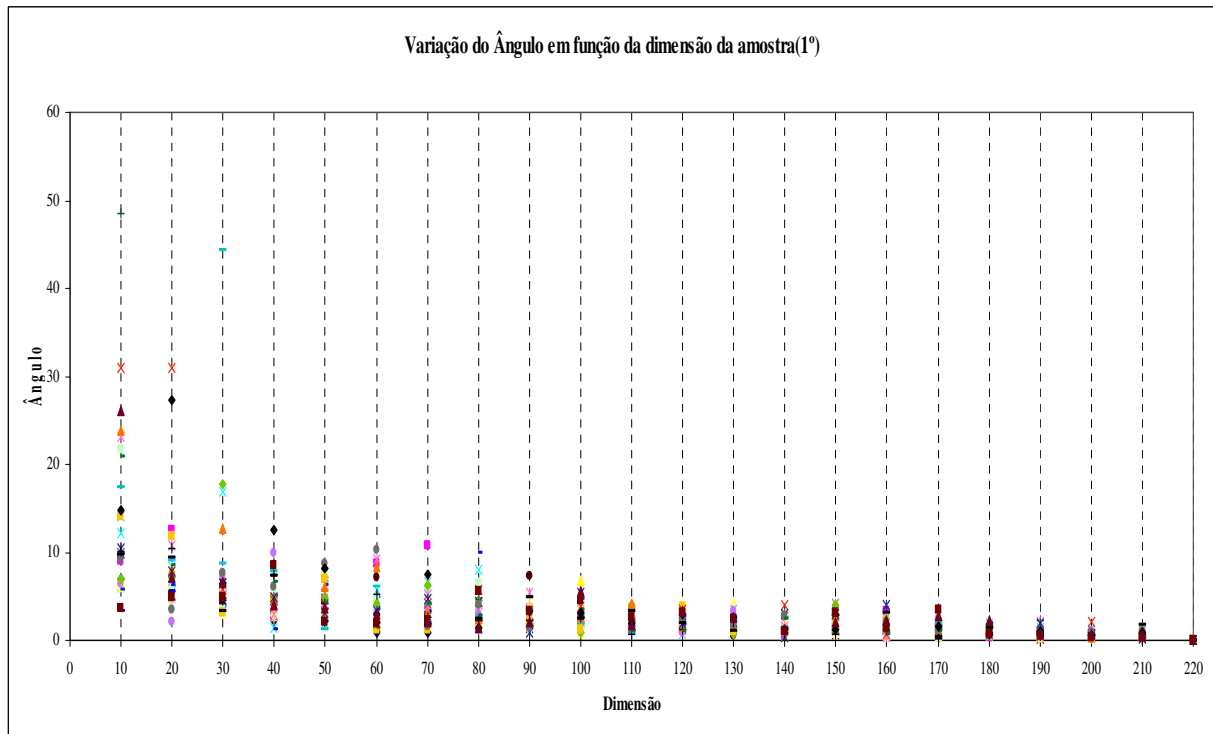


Fig.3.23 -Variação do primeiro ângulo em função da dimensão da amostra

Observa-se que para amostras de pequenas dimensões a variação do ângulo é muito grande. Para amostras de dimensão 10, por exemplo, o ângulo varia desde um máximo de $48,99^\circ$ até um mínimo de $3,37^\circ$.

No entanto, para maiores dimensões o valor do ângulo estabiliza, sendo muito reduzido. De facto, para amostras de dimensão 100 o ângulo máximo atingido é de $6,783^\circ$.

Pode concluir-se que a estabilidade do sistema aumenta com o aumento da dimensão da amostra, uma vez que se verifica que o ângulo vai diminuindo (sobre conceito de estabilidade ver Greenacre ([14])).

O ângulo ultrapassa apenas duas vezes os 45° , considerado em [14] como limite de estabilidade: para uma amostra de dimensão 10 e uma de dimensão 30.

O estudo do diagrama caixa de bigodes (Fig.3.24) mostra que, de facto, a variação do ângulo é maior para pequenas dimensões. Para amostras de dimensão 200 e 210 o diagrama revela a existência de outliers. No entanto, o valor máximo e mínimo atingido são, respectivamente, $2,0727^\circ$ e $0,2081^\circ$ para $n = 200$ e $1,6943^\circ$ e $0,0868^\circ$ para $n = 210$. São, evidentemente, valores muito diminutos.

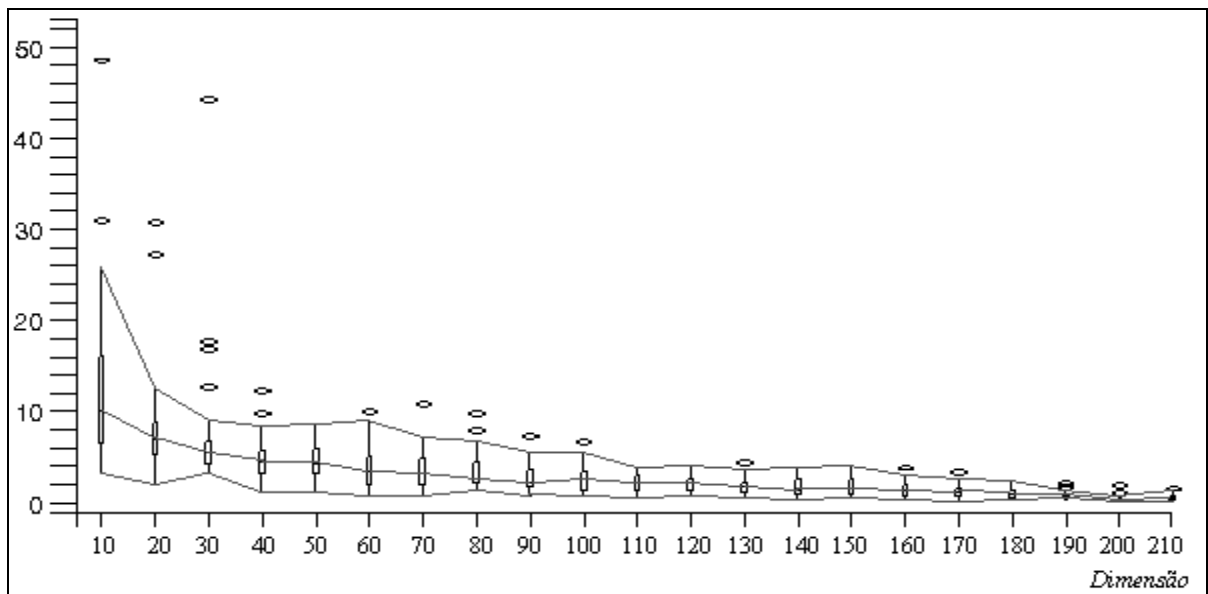


Fig.3.24 - Diagrama caixa de bigodes

3.5.2.2 Segunda Componente

Analisa-se de seguida a variação do ângulo entre o segundo eixo real e o ‘segundo eixo amostral’ (2° Ângulo) em função da dimensão da amostra (Fig.3.25).

Tal como acontece para o primeiro eixo, também a variação do ângulo entre o segundo eixo real e o segundo eixo ‘amostral’ é muito grande para amostras pequenas e vai diminuindo à medida que a dimensão aumenta (Fig.3.25).

Para $n = 10$, por exemplo, o ângulo varia entre $6,5^\circ$ e $51,3^\circ$, sendo que para amostras de dimensão superior a 100 o ângulo máximo atingido é sempre inferior a 6° .

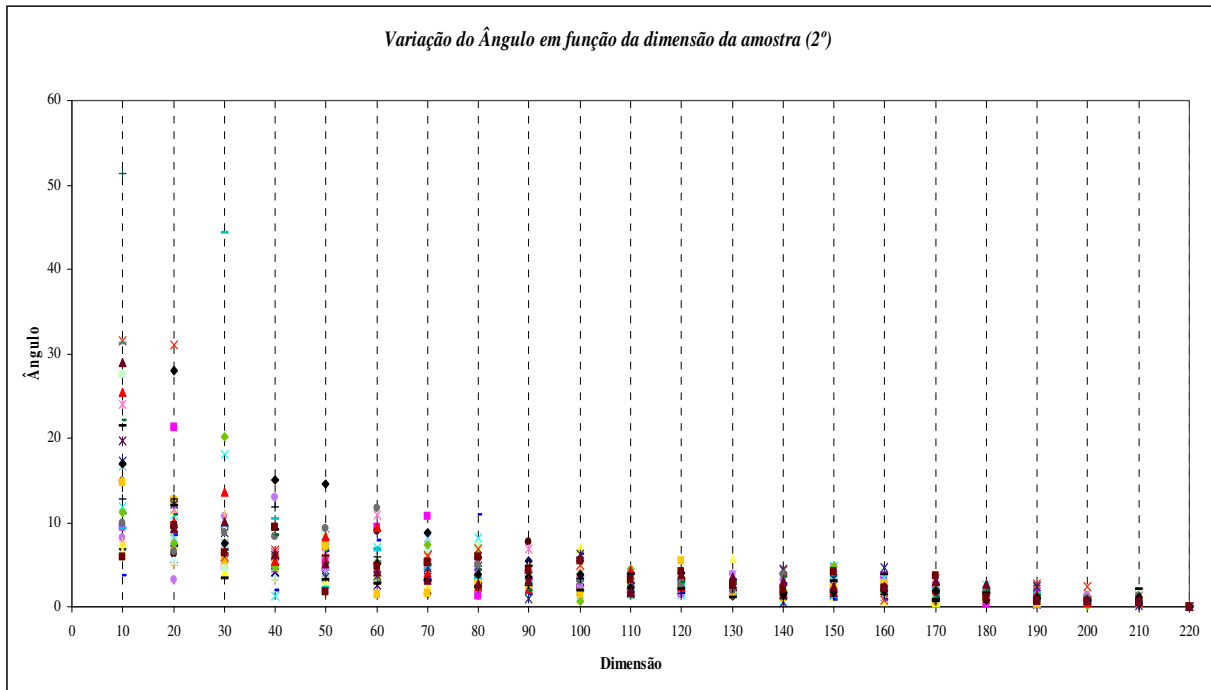


Fig.3.25 -Variação do segundo ângulo em função da dimensão da amostra

Para uma amostra de dimensão 10, o ângulo máximo obtido é superior a 51° , que ultrapassa o limite de estabilidade. Para $n = 30$ o ângulo obtido é muito próximo dos 45° : $44,36^\circ$. Portanto, o limite de estabilidade apenas é ultrapassado por uma vez.

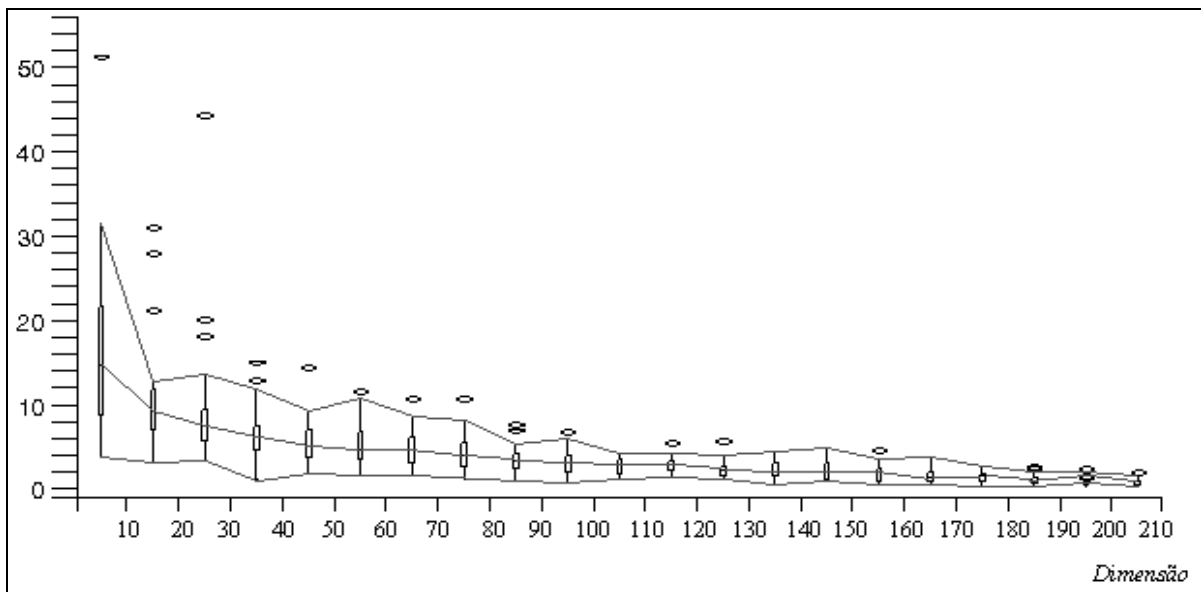


Fig.3.26 - Diagrama caixa de bigodes

3.5.2.3 Terceira Componente

Neste caso, ângulo entre o terceiro eixo de todos os paralelepípedos e o ‘terceiro eixo amostral’ (3° ângulo), o ângulo apenas estabiliza para amostras de grandes dimensões. Em particular, para amostras de dimensão superior a 150, a partir das quais o ângulo máximo atingido nunca é superior a 9° . Verifica-se também que em amostras de pequena dimensão a variação é maior que para o 1° e 2° vector próprio:

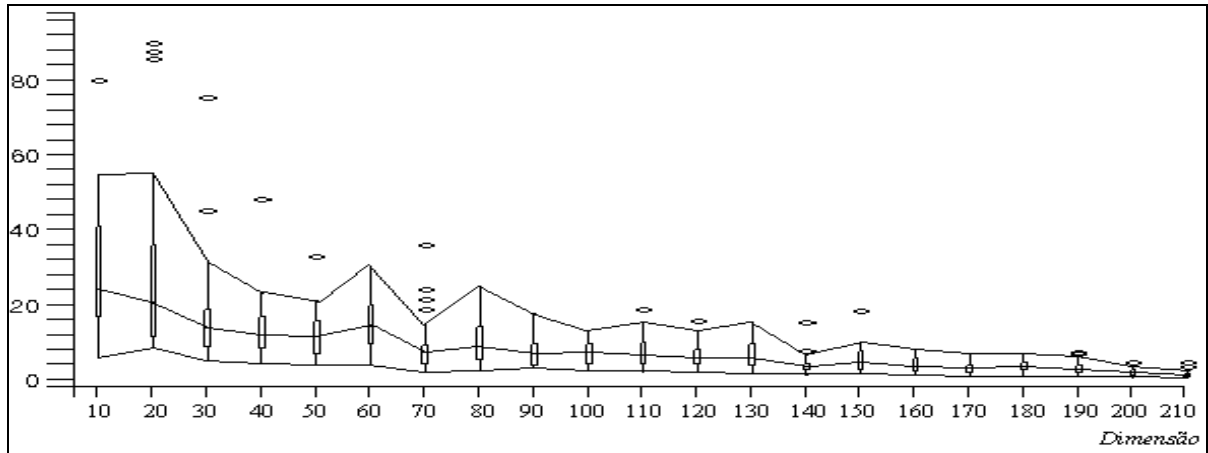


Fig.3.27 -Diagrama caixa de bigodes

Para amostras de dimensão 10, por exemplo, o ângulo varia desde um mínimo de $5,61^\circ$ até um máximo de $79,88^\circ$. Para amostras de dimensão 20, o ângulo máximo atinge $89,82^\circ$. Tem-se que para $n = 10$ o ângulo ultrapassa 4 vezes os 45° e atinge por vezes valores muito próximos dos 45° . Para $n = 20$ o limite de estabilidade é ultrapassado 5 vezes, 2 vezes para $n = 30$ e ainda uma vez para $n = 40$. Assim, o ângulo de variação inferior a 45° , considerado como limite de estabilidade, só é atingido para $n = 50$ (Fig.3.25):

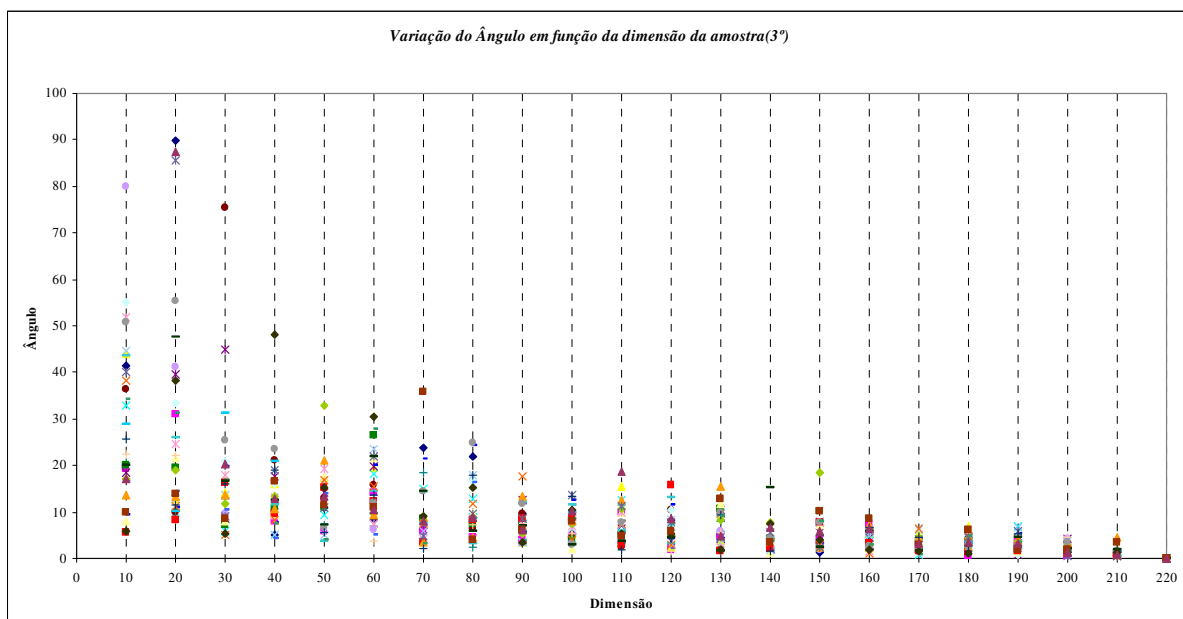


Fig.3.28 -Variação do terceiro ângulo em função da dimensão da amostra

A figura 3.29 representa a estabilidade comparada dos três ângulos.

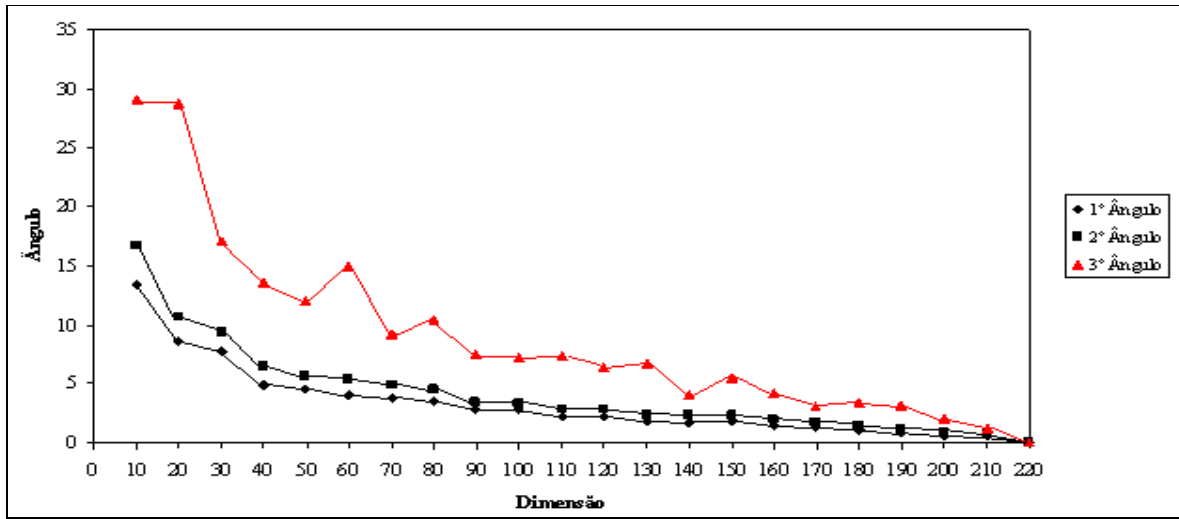


Fig.3.29 – Estabilidade Comparada

É claro que o primeiro e segundo ângulo estabilizam mais rapidamente que o terceiro ângulo. Além disso, os dois primeiros ângulos estabilizam monotonicamente o que não acontece com o terceiro.

Esta menor estabilidade deve-se, com certeza, ao facto de o terceiro eixo ser menos marcante em termos de inércia explicada.

Realce-se que combinando a estabilidade dos eixos se tem a estabilidade dos planos principais.

3.5.2.4 Análise da Normalidade do Ângulo

Tal como para o primeiro valor próprio a análise da normalidade do ângulo entre o primeiro eixo e o primeiro eixo amostral foi feita aplicando o teste de normalidade (γ_1, γ_2) ou de Kendall & Stuart.

Dimensão	Média	Desvio Padrão	Assimetria	Curtose
10	13,445	9,752	1,836	6,763
20	8,553	6,098	2,633	9,718
30	7,677	7,784	3,776	17,867
40	4,877	2,513	1,098	4,234
50	4,575	1,955	0,418	2,373
60	4,066	2,527	0,964	2,978
70	3,773	2,188	1,238	4,777
80	3,568	2,101	1,310	4,188
90	2,833	1,461	1,123	4,224
100	2,765	1,484	0,926	3,332
110	2,236	0,975	0,190	2,003
120	2,160	0,867	0,161	1,962
130	1,837	0,867	1,023	4,015
140	1,696	0,917	0,674	2,579
150	1,885	1,102	0,672	2,340
160	1,519	0,943	0,928	2,965
170	1,298	0,772	0,793	3,385
180	1,057	0,509	0,698	2,677
190	0,831	0,509	1,415	4,367
200	0,627	0,363	2,306	9,532
210	0,595	0,419	0,979	3,128

Tabela 3.11

Neste caso, a distribuição amostral do ângulo não passa testes de normalidade. Apenas para $n = 50, 110$ e 120 se verificam as condições de normalidade (Fig.3.30)

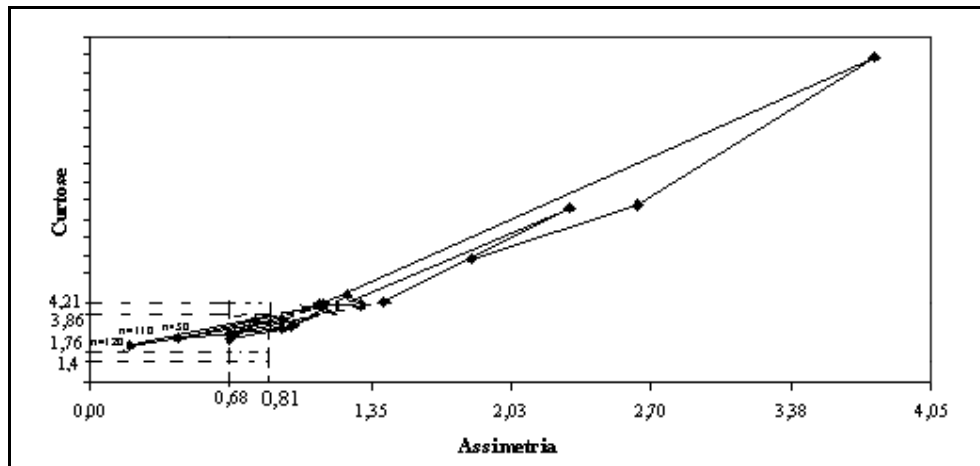


Fig.3.30- Representação Assimetria vs Curtose

Para ângulo entre o segundo eixo real e o segundo eixo apenas 33% das dimensões passam os testes de normalidade: $n = 100, 110, 120, 140, 150, 170$ e 180 .

Dimensão	Média	Desvio Padrão	Assimetria	Curtose
10	16,672	10,292	1,388	5,299
20	10,649	6,156	2,070	7,094
30	9,384	7,595	3,508	16,308
40	6,470	3,106	0,945	3,766
50	5,650	2,585	1,306	5,722
60	5,430	2,568	0,890	2,927
70	4,904	2,135	0,705	3,323
80	4,485	2,146	0,945	3,738
90	3,442	1,522	0,934	3,917
100	3,332	1,617	0,533	2,555
110	2,771	1,036	-0,065	1,780
120	2,847	0,943	0,431	3,225
130	2,327	0,975	1,510	5,872
140	2,292	1,154	0,438	2,094
150	2,368	1,205	0,658	2,368
160	1,966	1,028	0,701	2,992
170	1,667	0,882	0,589	2,361
180	1,483	0,590	0,406	2,602
190	1,128	0,691	1,158	3,060
200	0,803	0,432	1,980	7,840
210	0,738	0,454	1,066	3,480

Tabela 3.12

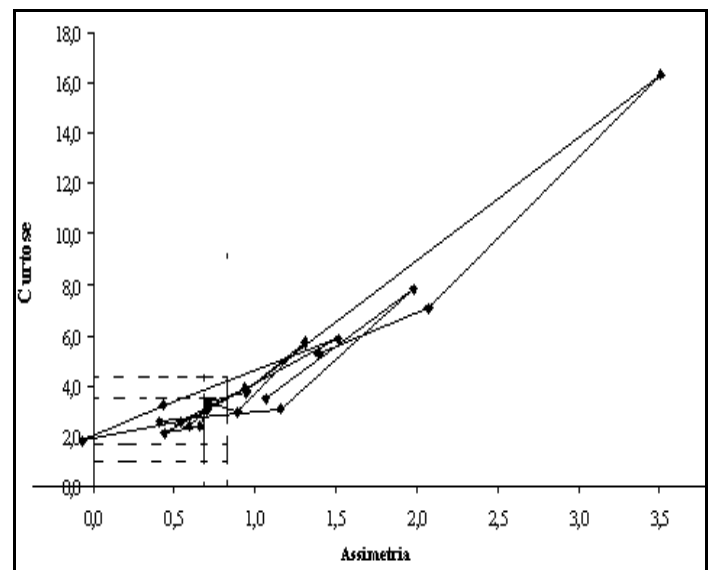


Fig.3.31 - Representação Assimetria vs Curtose

3.6 Comparação com Paralelepípedos de J. Davis

Aos paralelepípedos considerados em [1] (Anexo III), e apresentados no início deste capítulo, aplicaram-se as variáveis escolhidas para este estudo.

Realizando ACP obteve-se a representação dos indivíduos no plano formado pelos dois primeiros eixos principais de inércia que se apresenta na figura 3.32. Compara-se esta representação com a obtida para os mesmos paralelepípedos mas tomando as variáveis de [1] (Fig.3.33).

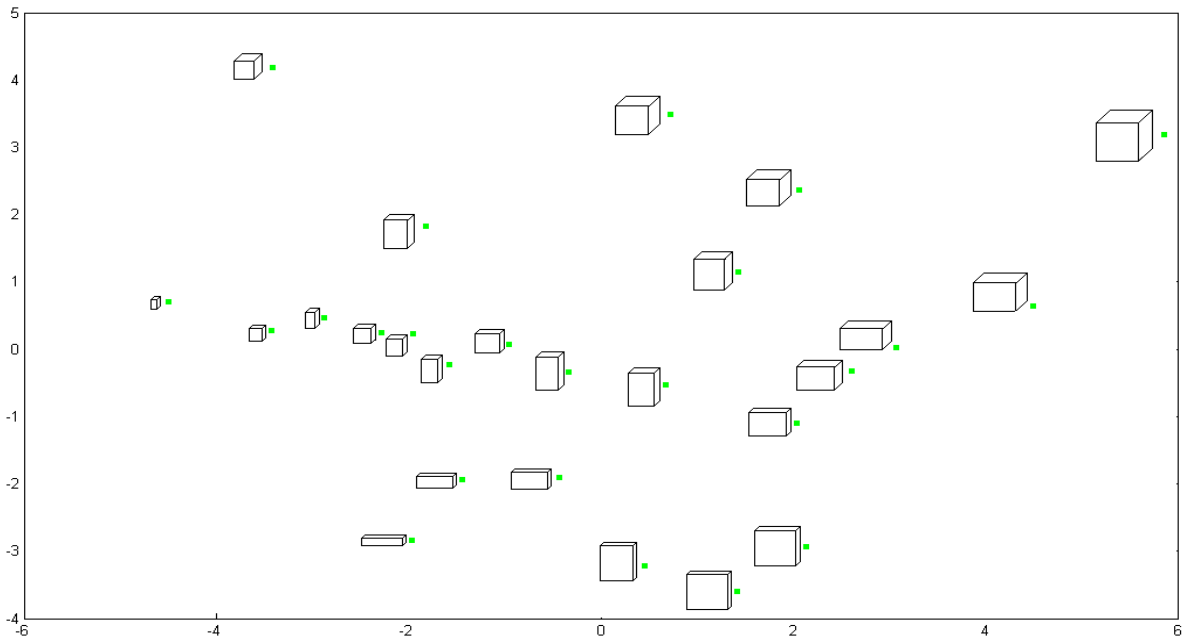


Fig.3.32 - Indivíduos representados no plano principal (1,2)

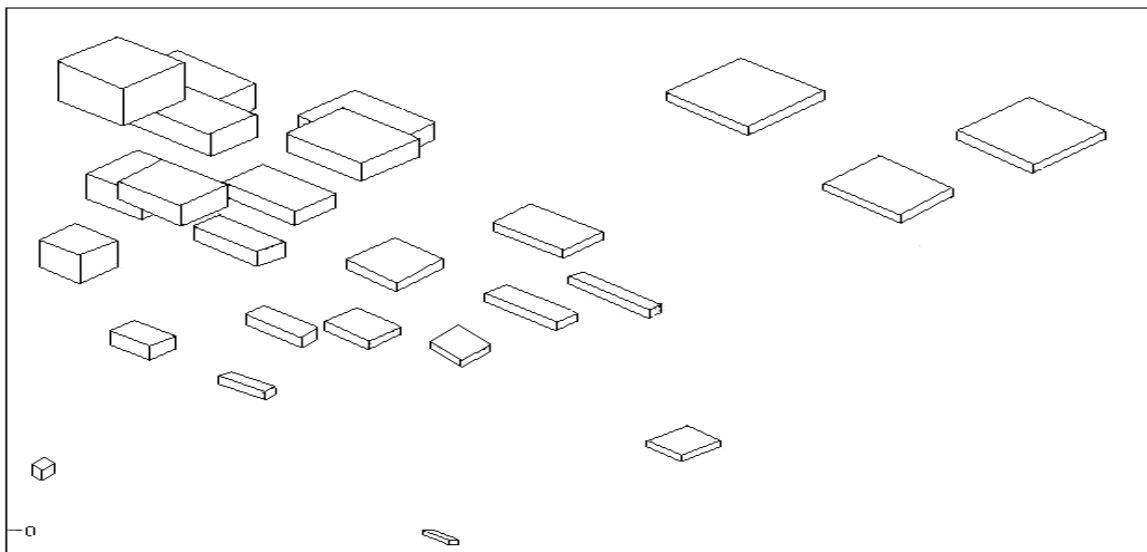


Fig.3.33 - Representação FC-FS

Observa-se que aplicando as variáveis deste estudo (Tabela 3.1), inclusive menos sofisticadas, a ACP permite uma melhor distinção entre os paralelepípedos. Quando se utilizam as variáveis de [1] permanece no centro do gráfico uma mistura entre *tábuas* e *palitos* (Fig.3.35):

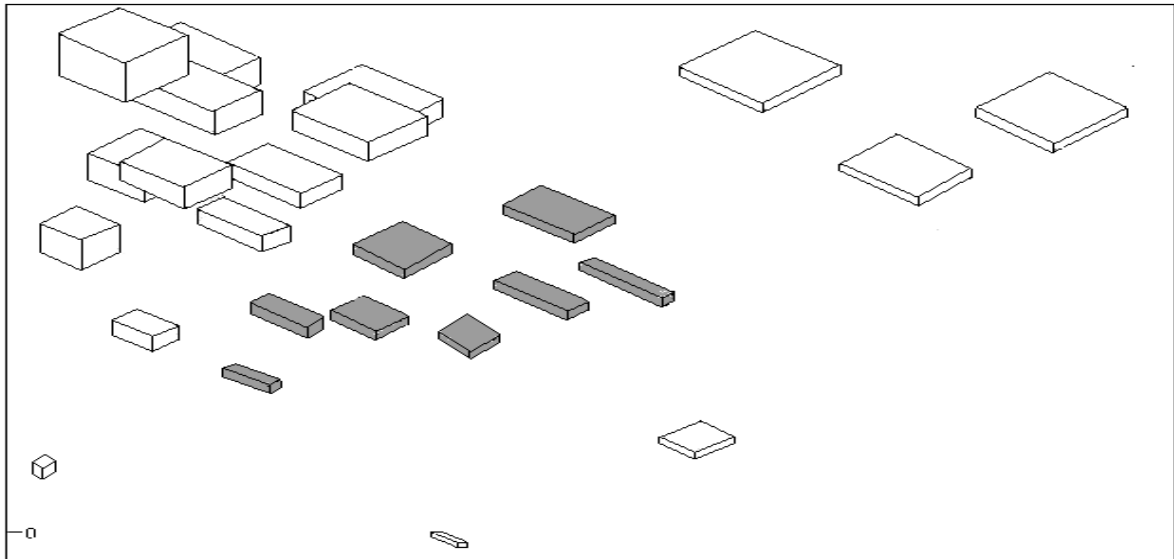


Fig.3.34 – Representação FC-FS

Para as variáveis consideradas, a organização dos paralelepípedos de [1] é idêntica à obtida para os paralelepípedos deste estudo: grandes cubos num vértice, pequenos cubos noutra vértice e tábuas e palitos também em vértices distintos.

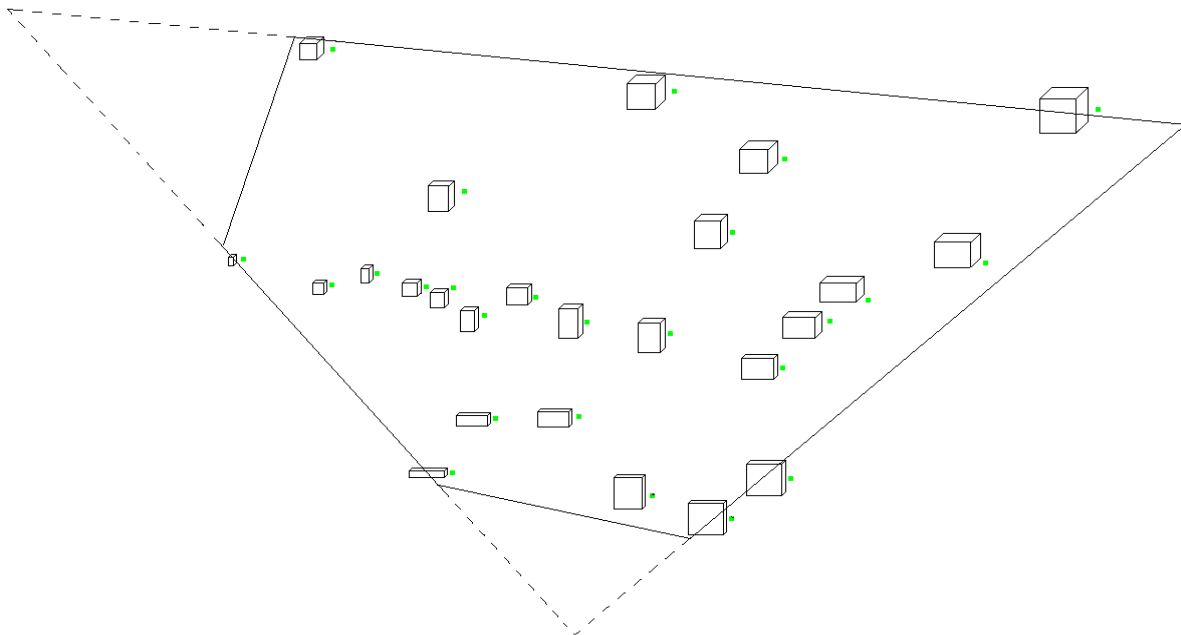


Fig.3.35 – Organização dos paralelepípedos FC-FS considerando as variáveis deste estudo

Capítulo IV

Variantes

Neste ponto, segue-se o estudo de algumas variantes do primeiro ficheiro: dimensões das arestas não ordenadas, arestas contínuas, classificações de alunos e, por fim, uma família multinormal com matriz de correlações igual à obtida na Tabela 3.3.

4.1 Dimensões não Ordenadas

Pretende-se analisar o efeito da ordenação na organização dos paralelepípedos quando é realizada uma ACP. Para tal efectuou-se o mesmo procedimento, mas não ordenando as dimensões.

Dada a não ordenação, deixa de ter sentido a igualdade $X_1 = X_7$ e $X_3 = X_6$.

Obteve-se a seguinte matriz de correlações:

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1	0,11	0,12	0,53	0,60	0,50	0,45	0,64	0,60	-0,15	0,04	0,18
X2	0,11	1	0,05	0,56	0,61	0,51	0,51	0,64	0,61	-0,21	0,00	0,25
X3	0,12	0,05	1	0,56	0,60	0,45	0,50	0,61	0,60	-0,18	-0,04	0,27
X4	0,53	0,56	0,56	1	0,97	0,50	0,91	0,88	0,97	-0,49	-0,35	0,61
X5	0,60	0,61	0,60	0,97	1	0,63	0,86	0,97	1,00	-0,42	-0,20	0,52
X6	0,50	0,51	0,45	0,50	0,63	1	0,33	0,77	0,63	0,21	0,58	-0,18
X7	0,45	0,51	0,50	0,91	0,86	0,33	1	0,77	0,86	-0,71	-0,58	0,83
X8	0,64	0,64	0,61	0,88	0,97	0,77	0,77	1	0,97	-0,29	0,00	0,37
X9	0,60	0,61	0,60	0,97	1	0,63	0,86	0,97	1	-0,42	-0,20	0,52
X10	-0,15	-0,21	-0,18	-0,49	-0,42	0,21	-0,71	-0,29	-0,42	1	0,79	-0,82
X11	0,04	0,00	-0,04	-0,35	-0,20	0,58	-0,58	0,00	-0,20	0,79	1	-0,87
X12	0,18	0,25	0,27	0,61	0,52	-0,18	0,83	0,37	0,52	-0,82	-0,87	1

Tabela 4.1

Procurou-se organizá-la em blocos de fortes correlações:

	X4	X5	X9	X8	X7	X10	X11	X12	X1	X2	X3	X6
X4	1	0,97	0,97	0,88	0,91	-0,49	-0,35	0,61	0,53	0,56	0,56	0,50
X5	0,97	1	1,00	0,97	0,86	-0,42	-0,20	0,52	0,60	0,61	0,60	0,63
X9	0,97	1	1	0,97	0,86	-0,42	-0,20	0,52	0,60	0,61	0,60	0,63
X8	0,88	0,97	0,97	1	0,77	-0,29	0,00	0,37	0,64	0,64	0,61	0,77
X7	0,91	0,86	0,86	0,77	1	-0,71	-0,58	0,83	0,45	0,51	0,50	0,33
X10	-0,49	-0,42	-0,42	-0,29	-0,71	1	0,79	-0,82	-0,15	-0,21	-0,18	0,21
X11	-0,35	-0,20	-0,20	0,00	-0,58	0,79	1	-0,87	0,04	0,00	-0,04	0,58
X12	0,61	0,52	0,52	0,37	0,83	-0,82	-0,87	1	0,18	0,25	0,27	-0,18
X1	0,53	0,60	0,60	0,64	0,45	-0,15	0,04	0,18	1	0,11	0,12	0,50
X2	0,56	0,61	0,61	0,64	0,51	-0,21	0,00	0,25	0,11	1	0,05	0,51
X3	0,56	0,60	0,60	0,61	0,50	-0,18	-0,04	0,27	0,12	0,05	1	0,45
X6	0,50	0,63	0,63	0,77	0,33	0,21	0,58	-0,18	0,50	0,51	0,45	1

Tabela 4.2

Continua a manter-se um bloco formado pelas variáveis analíticas (X_4, X_5, X_7, X_8, X_9), outro bloco formado pelas variáveis função dos extremos (X_{10}, X_{11}, X_{12}) e a variável máximo, X_6 , permanece isolada. No entanto, nesta situação as variáveis representativas das dimensões não podem ser incluídas em nenhum dos blocos. Apresentam correlação significativa com as variáveis analíticas, mas estão fracamente relacionadas entre si.

Na Tabela 4.3 estão apresentados os valores próprios obtidos para a matriz de correlações:

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
6,79	56,66	57
2,88	24,04	81
0,95	7,95	89
0,86	7,16	96
0,25	2,07	98
0,14	1,18	99
0,08	0,69	100
0,03	0,25	100
0,00	0,02	100
0,00	0,00	100
0,00	0,00	100
0,00	0,00	100

Tabela 4.3

Observa-se que desordenando as dimensões diminui o total da inércia explicada pelas duas primeiras componentes (80,7%), ganhando uma maior importância as 3ª e 4ª componentes.

No entanto, observando a representação dos indivíduos no plano principal (1,2) verifica-se que não existem alterações significativas a apontar.

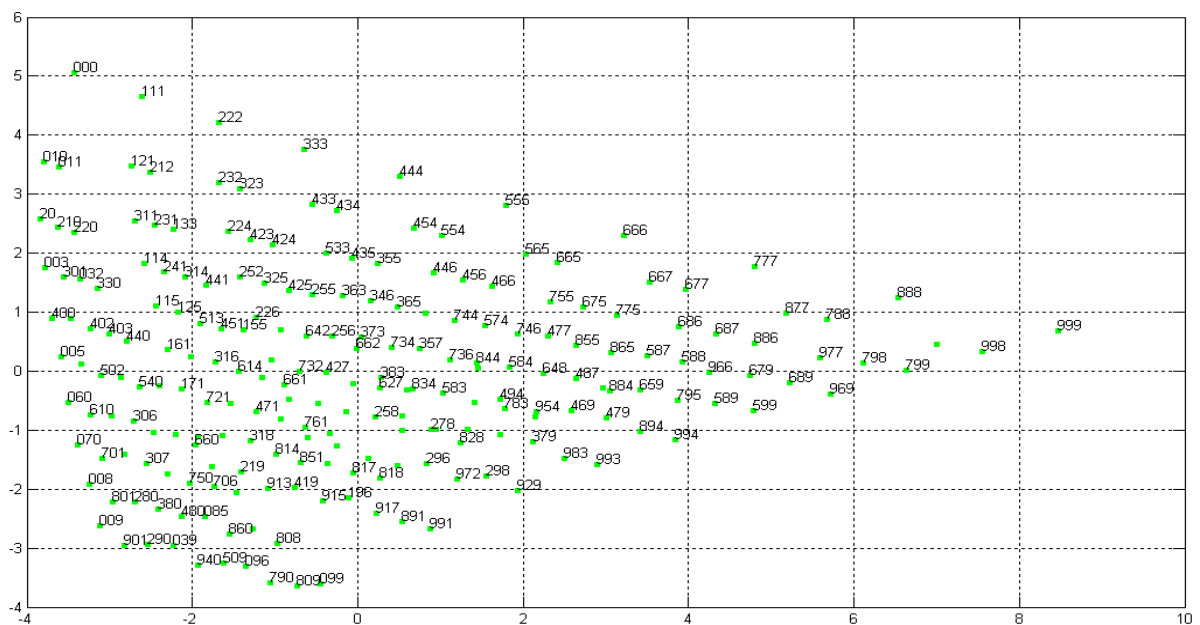


Fig.4.1 – Representação dos Indivíduos no plano principal (1,2)

A estrutura triangular ou trapezoidal mantém-se, bem como a interpretação de cada aresta.

Aliás, em termos de círculo de correlações, o único ponto a assinalar é o facto de, agora, todas as dimensões se encontrarem no interior do círculo. Portanto, as variáveis mais importantes para a formação das primeiras componentes mantêm-se inalteradas.

	Y1	Y2
X1	0,57	-0,28
X2	0,60	-0,20
X3	0,59	-0,16
X4	0,97	0,00
X5	0,98	-0,16
X6	0,55	-0,80
X7	0,94	0,27
X8	0,93	-0,34
X9	0,98	-0,16
X10	-0,56	-0,71
X11	-0,34	-0,93
X12	0,66	0,71

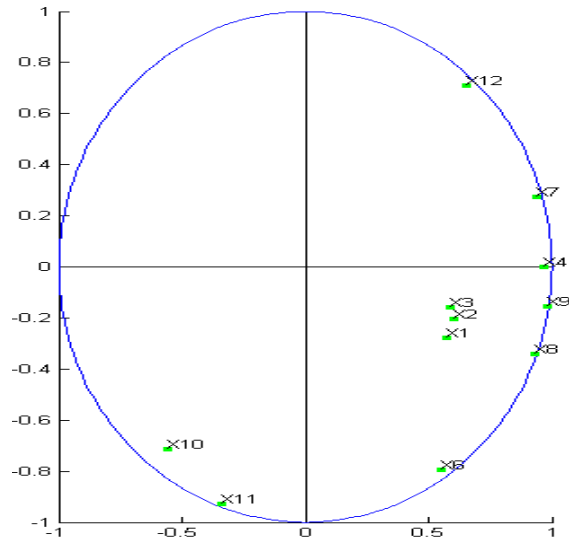


Fig.4.2 – Círculo de Correlações (12)

Quanto ao terceiro eixo já se podem observar alterações. A estrutura obtida considerando as dimensões ordenadas apresenta-se atenuada:

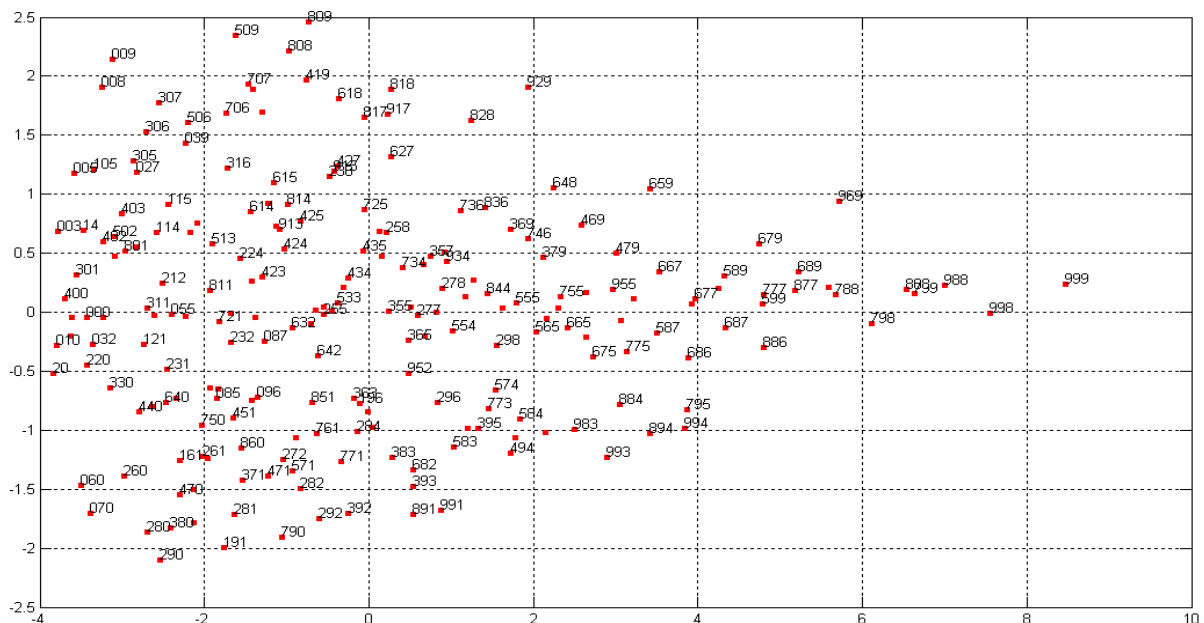


Fig.4.3 – Representação dos Indivíduos no plano principal (1,3)

Este eixo parecia, no caso ordenado, distinguir paralelepípedos tipo ‘tábua’ e tipo ‘palito’, sendo a variável X_2 a que mais contribuía para a formação da segunda componente.

Nesta situação, tem-se a variável X_2 que apresenta maior contributo, mas também a variável X_3 apresenta contributo importante.

	Y3
X1	0,09
X2	-0,71
X3	0,66
X4	0,01
X5	0,01
X6	-0,02
X7	-0,01
X8	0,00
X9	0,01
X10	0,05
X11	-0,01
X12	0,00

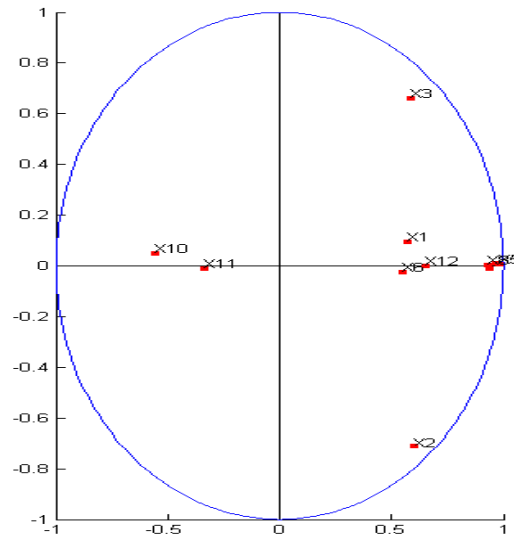


Fig. 4.4 – Círculo de Correlações (13)

A distinção que parece ser feita por este eixo ocorre entre paralelepípedos com segunda dimensão pequena e terceira dimensão grande e paralelepípedos em que ocorre o contrário.

Uma vez que a 4ª componente explica a mesma quantidade de informação que a 3ª analisa-se a sua importância em termos de interpretação.

X_1 é a variável que mais contribui para a formação desta componente, observando-se que no 4º eixo se opõem paralelepípedos com valor elevado e reduzido nesta dimensão:

	Y4
X1	0,76
X2	-0,30
X3	-0,43
X4	-0,01
X5	0,00
X6	-0,04
X7	-0,01
X8	0,00
X9	0,00
X10	-0,04
X11	-0,03
X12	0,00

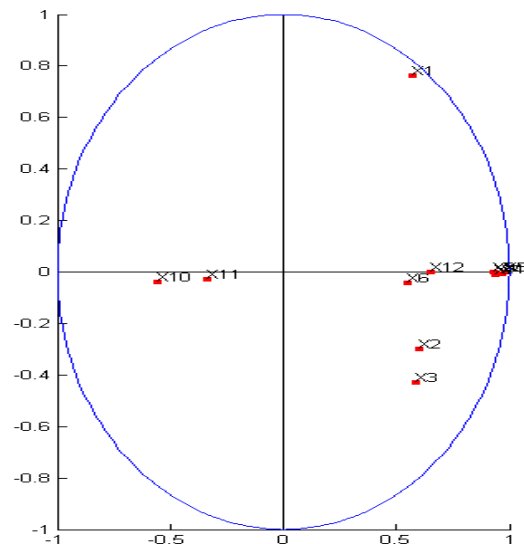


Fig. 4.5 – Círculo de Correlações (14)

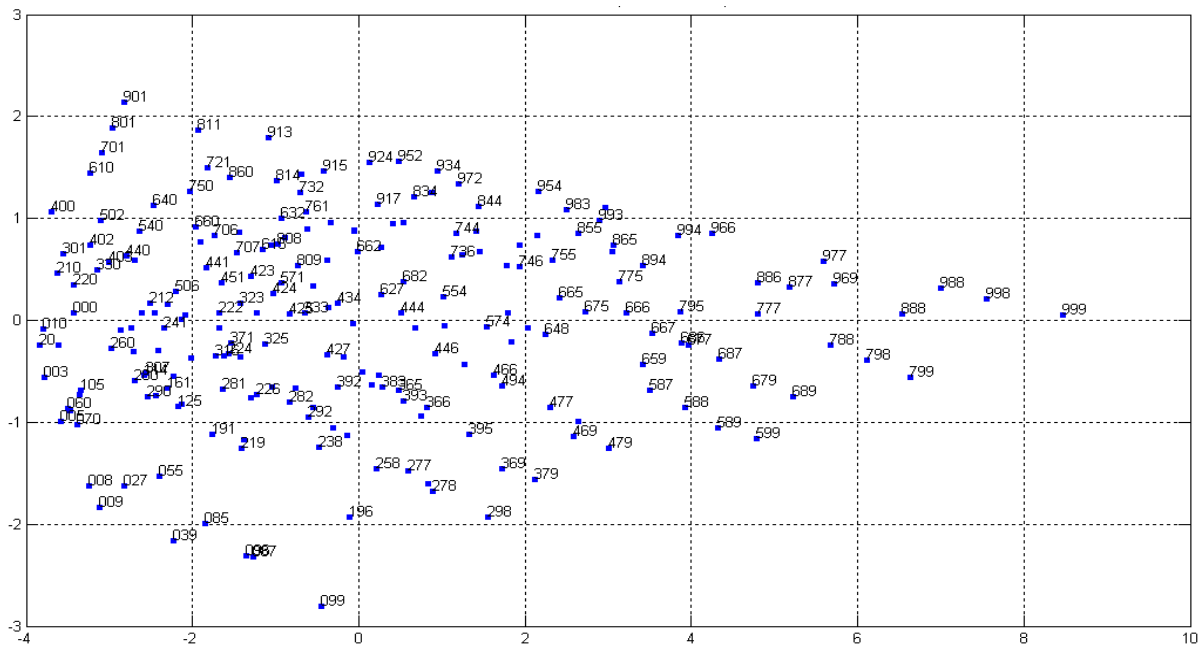


Fig. 4.6 – Representação dos Indivíduos no plano principal (1,4)

Notoriamente, em termos de estrutura, as figuras 4.3 e 4.6 são muito semelhantes.

A ordenação não produz efeitos significativos nas duas primeiras componentes, que continuam a ser suficientes para o resumo da informação total

A alteração mais importante ocorre na terceira componente que passa a ter uma nova interpretação.

4.2 Hiper-Paralelepípedos de Dimensão Quatro

Analisa-se hiper-paralelepípedos de dimensão quatro, ou seja, constituídos por 2^4 vértices.

De entre todos os paralelepípedos possíveis foram escolhidos, ao acaso, 220, para termos um número igual ao total do caso tridimensional (Anexo V).

Para estes indivíduos considerou-se um conjunto de 25 variáveis, análogas às anteriormente consideradas:

X_1	<i>Comprimento da Aresta</i>	→	a
X_2	<i>Comprimento da Aresta</i>	→	b
X_3	<i>Comprimento da Aresta</i>	→	c
X_4	<i>Comprimento da Aresta</i>	→	d
X_5	<i>HiperVolume</i>	→	$X_1 X_2 X_3 X_4$
X_6	<i>‘Soma de Bi-Produtos’</i>	→	$X_1 X_2 + X_1 X_3 + X_1 X_4 + X_2 X_3 + X_2 X_4 + X_3 X_4$
X_7	<i>‘Soma de Tri-Produtos’</i>	→	$X_1 X_2 X_3 + X_1 X_2 X_4 + X_2 X_3 X_4 + X_1 X_3 X_4$
X_8	<i>Dimensão Máxima</i>	→	$Max(X_1, X_2, X_3, X_4)$
X_9	<i>Dimensão Mínima</i>	→	$Min(X_1, X_2, X_3, X_4)$
X_{10}	<i>Média das dimensões</i>	→	$Media(X_1, X_2, X_3, X_4)$
X_{11}	<i>Média dos Bi-‘Produtos’</i>	→	$Media(X_1 X_2, X_1 X_3, X_1 X_4, X_2 X_3, X_2 X_4, X_3 X_4)$
X_{12}	<i>Média dos Tri-‘Produtos’</i>	→	$Media(X_1 X_2 X_3, X_1 X_2 X_4, X_2 X_3 X_4, X_1 X_3 X_4)$
X_{13}	<i>Quocientes entre Dimensões</i>	→	$\frac{X_4}{X_3}$
X_{14}		→	$\frac{X_4}{X_2}$
X_{15}		→	$\frac{X_4}{X_1}$
X_{16}		→	$\frac{X_3}{X_2}$
X_{17}		→	$\frac{X_3}{X_1}$
X_{18}		→	$\frac{X_2}{X_1}$
X_{19}	<i>Diferenças entre Dimensões</i>	→	$X_4 - X_3$
X_{20}		→	$X_4 - X_2$
X_{21}		→	$X_4 - X_1$
X_{22}		→	$X_3 - X_2$
X_{23}		→	$X_3 - X_1$
X_{24}		→	$X_2 - X_1$
X_{25}	<i>Mínimo áreas laterais</i> <i>Máximo áreas laterais</i>	→	$\frac{Min(X_1 X_2 X_3, X_1 X_2 X_4, X_2 X_3 X_4, X_1 X_3 X_4)}{Max(X_1 X_2 X_3, X_1 X_2 X_4, X_2 X_3 X_4, X_1 X_3 X_4)}$

Tabela 4.4

Como se consideram novamente as dimensões ordenadas, tem-se que $X_1 = X_9$ e $X_4 = X_8$.

Dada a similaridade das variáveis com as consideradas a três dimensões, repetem-se algumas relações detectadas anteriormente. Em particular, tem-se que:

$$X_{11} = \text{Media}(X_1X_2, X_1X_3, X_1X_4, X_2X_3, X_2X_4, X_3X_4) = \frac{X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4}{6} = \frac{X_6}{6};$$

$$X_{12} = \text{Media}(X_1X_2X_3, X_1X_2X_4, X_2X_3X_4, X_1X_3X_4) = \frac{X_1X_2X_3 + X_1X_2X_4 + X_2X_3X_4 + X_1X_3X_4}{4} = \frac{X_7}{4}$$

$$\text{e } X_{25} = \frac{\text{Min}(X_1X_2X_3, X_1X_2X_4, X_2X_3X_4, X_1X_3X_4)}{\text{Max}(X_1X_2X_3, X_1X_2X_4, X_2X_3X_4, X_1X_3X_4)} = \frac{X_1X_2X_3}{X_2X_3X_4} = \frac{X_1}{X_4} = \frac{1}{X_{15}}.$$

Procurou-se organizar a matriz de correlações obtida (Tabela 4.5) em blocos fortemente correlacionados (Tabela 4.6). Nestes blocos todas as correlações são superiores, em módulo, a 50%. Existem, naturalmente, algumas correlações fortes exteriores aos blocos, que se encontram assinaladas. Aliás, neste caso, a organização em blocos torna-se mais complicada, existindo várias possibilidades para a definição dos blocos. Tentou-se, contudo, que os blocos incluíssem as maiores correlações possíveis.

Verifica-se que as variáveis $X_5, X_6, X_7, X_9, X_{10}, X_{11}$ e X_{12} se encontram fortemente correlacionadas: todas elas são funções analíticas (polinómios em várias variáveis). Deste grupo, fazem também parte as duas menores dimensões, X_1 e X_2 . Constituem o bloco de maior dimensão.

Num segundo bloco distinguem-se variáveis que representam quocientes e diferenças entre as duas maiores dimensões e a dimensão mínima: $X_{15}, X_{21}, X_{25}, X_{17}$ e X_{23} .

Diferenças e quocientes entre as duas maiores dimensões, X_3 e X_4 , e X_2 formam o terceiro maior bloco de fortes correlações: X_{14}, X_{16}, X_{20} e X_{22} .

X_{18} e X_{24} incluem-se num outro bloco: representam a diferença e o quociente entre as menores dimensões.

A segunda maior dimensão e as suas relações com X_4 (X_{13} e X_{19}) compõem também um bloco.

Finalmente, tem-se, isolada, a dimensão máxima, X_4 ($= X_8$).

Em relação às correlações fortes existentes nas antidiagonais pode destacar-se a correlação entre X_1 as variáveis X_{15}, X_{17}, X_{21} e X_{25} . Também as variáveis X_{23} e X_{17} estão muito correlacionadas com as variáveis X_{18} e X_{24} .

Têm-se também, evidentemente, que X_1 , X_2 e X_3 se correlacionam fortemente com médias e produtos.

Para dimensão três podiam, aparentemente, distinguir-se três blocos: variáveis analíticas, variáveis função dos extremos e dimensão máxima. Para a dimensão agora estudada a organização das correlações

é idêntica. No entanto, acrescentam-se novos blocos fruto do maior número de relações que é possível estabelecer entre as dimensões.

Sinteticamente podem considerar-se três tipos de blocos: variáveis analíticas, variáveis função das dimensões (amplitudes) e máximo.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25
X1	1																								
X2	0,65	1																							
X3	0,37	0,68	1																						
X4	0,20	0,33	0,63	1																					
X5	0,85	0,76	0,57	0,36	1																				
X6	0,78	0,89	0,80	0,56	0,91	1																			
X7	0,84	0,85	0,69	0,45	0,98	0,98	1																		
X8	0,20	0,33	0,63	1	0,36	0,56	0,45	1																	
X9	1	0,65	0,37	0,20	0,85	0,78	0,84	0,20	1																
X10	0,70	0,86	0,87	0,68	0,81	0,97	0,91	0,68	0,70	1															
X11	0,78	0,89	0,80	0,56	0,91	1	0,98	0,56	0,78	0,97	1														
X12	0,84	0,85	0,69	0,45	0,98	0,98	1	0,45	0,84	0,91	0,98	1													
X13	-0,24	-0,42	-0,60	0,01	-0,26	-0,40	-0,33	0,01	-0,24	-0,42	-0,40	-0,33	1												
X14	-0,49	-0,72	-0,30	0,12	-0,42	-0,51	-0,49	0,12	-0,49	-0,46	-0,51	-0,49	0,48	1											
X15	-0,75	-0,39	-0,01	0,25	-0,47	-0,39	-0,45	0,25	-0,75	-0,28	-0,39	-0,45	0,22	0,60	1										
X16	-0,39	-0,53	0,11	0,15	-0,29	-0,29	-0,32	0,15	-0,39	-0,22	-0,29	-0,32	-0,09	0,80	0,52	1									
X17	-0,61	-0,12	0,34	0,25	-0,32	-0,14	-0,25	0,25	-0,61	-0,03	-0,14	-0,25	-0,22	0,31	0,86	0,53	1								
X18	-0,39	0,36	0,31	0,15	-0,12	0,09	0,00	0,15	-0,39	0,16	0,09	0,00	-0,19	-0,26	0,54	-0,17	0,70	1							
X19	-0,24	-0,47	-0,54	0,32	-0,30	-0,37	-0,35	0,32	-0,24	-0,32	-0,37	-0,35	0,75	0,50	0,28	0,02	-0,15	-0,22	1						
X20	-0,44	-0,65	-0,12	0,50	-0,40	-0,36	-0,42	0,50	-0,44	-0,24	-0,36	-0,42	0,39	0,76	0,56	0,60	0,30	-0,21	0,69	1					
X21	-0,62	-0,25	0,21	0,64	-0,38	-0,17	-0,29	0,64	-0,62	0,00	-0,17	-0,29	0,20	0,49	0,78	0,42	0,68	0,42	0,44	0,74	1				
X22	-0,36	-0,42	0,38	0,37	-0,25	-0,13	-0,22	0,37	-0,36	-0,01	-0,13	-0,22	-0,22	0,54	0,48	0,81	0,57	-0,06	-0,07	0,68	0,58	1			
X23	-0,47	0,10	0,64	0,43	-0,16	0,11	-0,03	0,43	-0,47	0,24	0,11	-0,03	-0,37	0,12	0,61	0,43	0,83	0,62	-0,31	0,25	0,72	0,66	1		
X24	-0,26	0,56	0,46	0,20	0,04	0,28	0,18	0,20	-0,26	0,33	0,28	0,18	-0,27	-0,39	0,32	-0,26	0,52	0,87	-0,35	-0,36	0,36	-0,14	0,65	1	
X25	0,85	0,48	0,07	-0,28	0,60	0,47	0,55	-0,28	0,85	0,35	0,47	0,55	-0,26	-0,55	-0,83	-0,44	-0,69	-0,42	-0,39	-0,66	-0,89	-0,51	-0,64	-0,32	1

Tabela 4.5 – Matriz de Correlações

	X12	X7	X5	X6	X11	X10	X2	X9	X1	X21	X15	X25	X17	X23	X18	X24	X3	X13	X19	X14	X16	X20	X22	X4	X8
X12	1																								
X7	1	1																							
X5	0,98	0,98	1																						
X6	0,98	0,98	0,91	1																					
X11	0,98	0,98	0,91	1	1																				
X10	0,91	0,91	0,81	0,97	0,97	1																			
X2	0,85	0,85	0,76	0,89	0,89	0,86	1																		
X9	0,84	0,84	0,85	0,78	0,78	0,7	0,65	1																	
X1	0,84	0,84	0,85	0,78	0,78	0,7	0,65	1	1																
X21	-0,29	-0,29	-0,38	-0,17	-0,17	0	-0,25	-0,62	-0,62	1															
X15	-0,45	-0,45	-0,47	-0,39	-0,39	-0,28	-0,39	-0,75	-0,75	0,78	1														
X25	0,55	0,55	0,6	0,47	0,47	0,35	0,48	0,85	0,85	-0,89	-0,83	1													
X17	-0,25	-0,25	-0,32	-0,14	-0,14	-0,03	-0,12	-0,61	-0,61	0,68	0,86	-0,69	1												
X23	-0,03	-0,03	-0,16	0,11	0,11	0,24	0,1	-0,47	-0,47	0,72	0,61	-0,64	0,83	1											
X18	0	0	-0,12	0,09	0,09	0,16	0,36	-0,39	-0,39	0,42	0,54	-0,42	0,7	0,62	1										
X24	0,18	0,18	0,04	0,28	0,28	0,33	0,56	-0,26	-0,26	0,36	0,32	-0,32	0,52	0,65	0,87	1									
X3	0,69	0,69	0,57	0,8	0,8	0,87	0,68	0,37	0,37	0,21	-0,01	0,07	0,34	0,64	0,31	0,46	1								
X13	-0,33	-0,33	-0,26	-0,4	-0,4	-0,42	-0,42	-0,24	-0,24	0,2	0,22	-0,26	-0,22	-0,37	-0,19	-0,27	-0,6	1							
X19	-0,35	-0,35	-0,3	-0,37	-0,37	-0,32	-0,47	-0,24	-0,24	0,44	0,28	-0,39	-0,15	-0,31	-0,22	-0,35	-0,54	0,75	1						
X14	-0,49	-0,49	-0,42	-0,51	-0,51	-0,46	-0,72	-0,49	-0,49	0,49	0,6	-0,55	0,31	0,12	-0,26	-0,39	-0,3	0,48	0,5	1					
X16	-0,32	-0,32	-0,29	-0,29	-0,29	-0,22	-0,53	-0,39	-0,39	0,42	0,52	-0,44	0,53	0,43	-0,17	-0,26	0,11	-0,09	0,02	0,8	1				
X20	-0,42	-0,42	-0,4	-0,36	-0,36	-0,24	-0,65	-0,44	-0,44	0,74	0,56	-0,66	0,3	0,25	-0,21	-0,36	-0,12	0,39	0,69	0,76	0,6	1			
X22	-0,22	-0,22	-0,25	-0,13	-0,13	-0,01	-0,42	-0,36	-0,36	0,58	0,48	-0,51	0,57	0,66	-0,06	-0,14	0,38	-0,22	-0,07	0,54	0,81	0,68	1		
X4	0,45	0,45	0,36	0,56	0,56	0,68	0,33	0,2	0,2	0,64	0,25	-0,28	0,25	0,43	0,15	0,2	0,63	0,01	0,32	0,12	0,15	0,5	0,37	1	
X8	0,45	0,45	0,36	0,56	0,56	0,68	0,33	0,2	0,2	0,64	0,25	-0,28	0,25	0,43	0,15	0,2	0,63	0,01	0,32	0,12	0,15	0,5	0,37	1	1

Tabela 4.6 – Matriz de Correlações Organizada

Verifica-se que as três primeiras componentes explicam 71% da inércia total, enquanto no caso tridimensional as duas primeiras componentes explicam 93% da informação:

[illegible]

Tabela 4.7

Uma explicação possível prende-se com a existência de um maior número de variáveis menos correlacionadas, o que significa uma maior quantidade de informação, mas mais dispersa por várias variáveis.

Observando a representação dos paralelepípedos na primeiro plano principal (Fig.4.7) verifica-se que os indivíduos se parecem organizar segundo um pentágono (Fig.4.8). No fundo, mantém-se a estrutura triangular, mas com os dois vértices menos agudos incompletos.

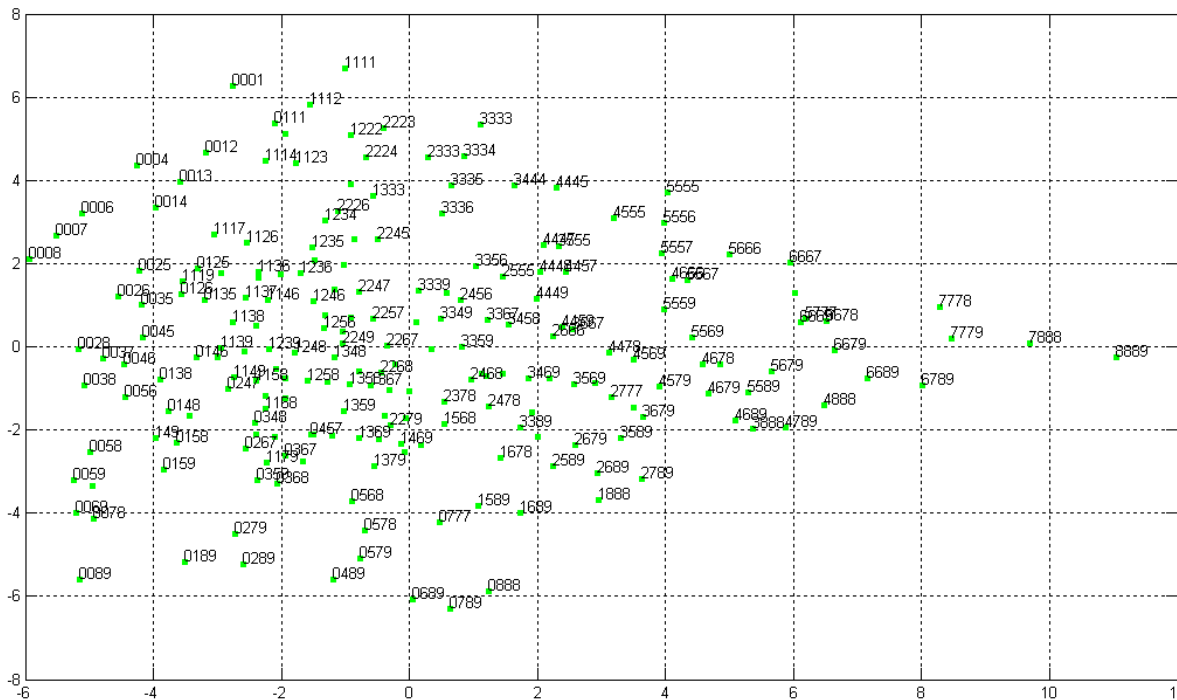


Fig.4.7 – Representação dos indivíduos no plano principal (1,2)

Na figura 4.8, nos vértices A e B encontram-se, respectivamente, os grandes e pequenos hipercubos. Na aresta [BC] observam-se paralelepípedos com três dimensões pequenas e iguais, sendo que próximo do vértice C se distinguem hiper palitos puros: 3 dimensões pequenas e uma grande. No vértice D têm-se paralelepípedos caracterizados por duas dimensões pequenas e duas grandes. Ao longo da aresta [DE] vai aumentando a segunda dimensão, pelo que em E se encontram as hiper tábuas puras.

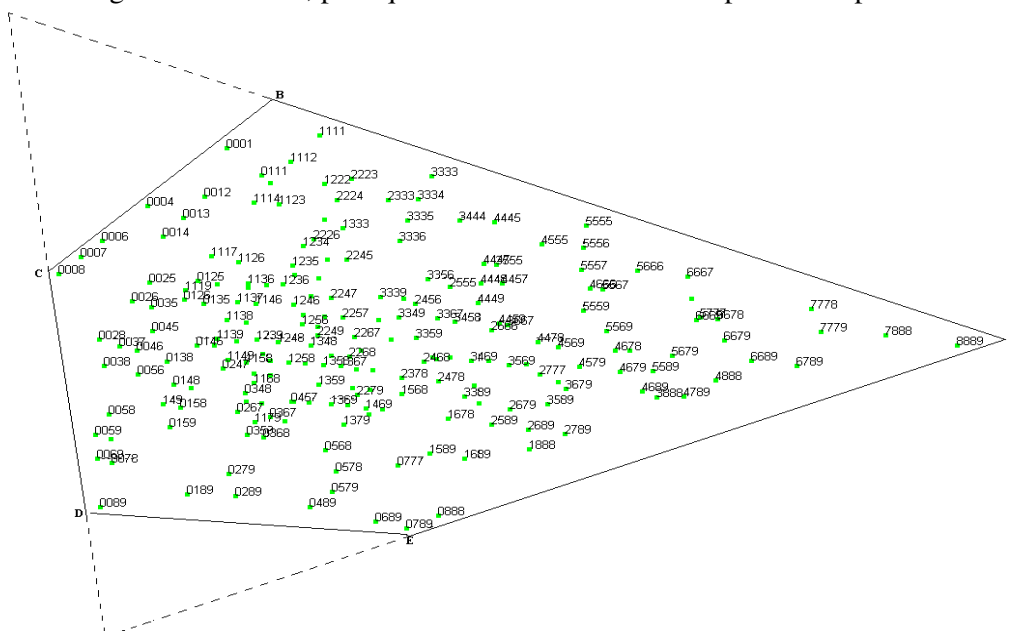


Fig.4.8 – Organização dos indivíduos

Interpretação das Componentes Principais

Neste caso, o primeiro eixo parece opor grandes e pequenos hipervolumes, hiperáreas e também biprodutos e triprodutos. De facto, encontram-se em oposição paralelepípedos caracterizados por todas as dimensões grandes (ou seja, grande hipervolume/ hiperárea /biproduto/triproduto) a paralelepípedos com pelo menos duas dimensões mínimas (e, portanto, pequeno hipervolume/ hiperárea/biproduto/triproduto).

Segue-se o círculo de correlações onde se podem distinguir 4 grupos de variáveis:

$\{X_1, X_2, X_5, X_6, X_7, X_9, X_{10}, X_{11}, X_{12}\}$ muito correlacionadas (positivamente) com a primeira componente, $\{X_4, X_8, X_{13}, X_{17}, X_{18}, X_{19}, X_{22}, X_{23}, X_{24}\}$ pouco correlacionadas e com correlações intermédias $\{X_{14}, X_{15}, X_{16}, X_{20}, X_{21}\}$ (negativamente) e $\{X_3, X_{25}\}$ (positivamente) :

	Y1
X1	0,9084
X2	0,8905
X3	0,5850
X4	0,2373
X5	0,9043
X6	0,9173
X7	0,9369
X8	0,2373
X9	0,9084
X10	0,8405
X11	0,9173
X12	0,9369
X13	-0,4173
X14	-0,7011
X15	-0,6708
X16	-0,5072
X17	-0,4221
X18	-0,0518
X19	-0,4567
X20	-0,6247
X21	-0,5198
X22	-0,3970
X23	-0,1939
X24	0,1470
X25	0,7493

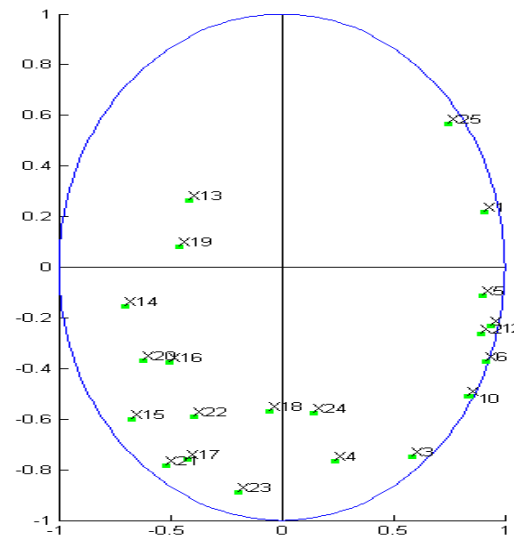


Fig.4.9 – Círculo de Correlações (12)

Ora, tem-se que o grupo de variáveis mais correlacionado com a primeira componente corresponde ao maior bloco construído na matriz de correlações. As variáveis com correlações intermédias pertencem aos blocos seguintes em termos de dimensão.

Seguem-se as linhas de nível para algumas das variáveis dominantes, em que o valor de cada linha aumenta no sentido positivo do eixo. As linhas equi-hipervolume são análogas às linhas equi-volume, embora nestas exista maior verticalidade.

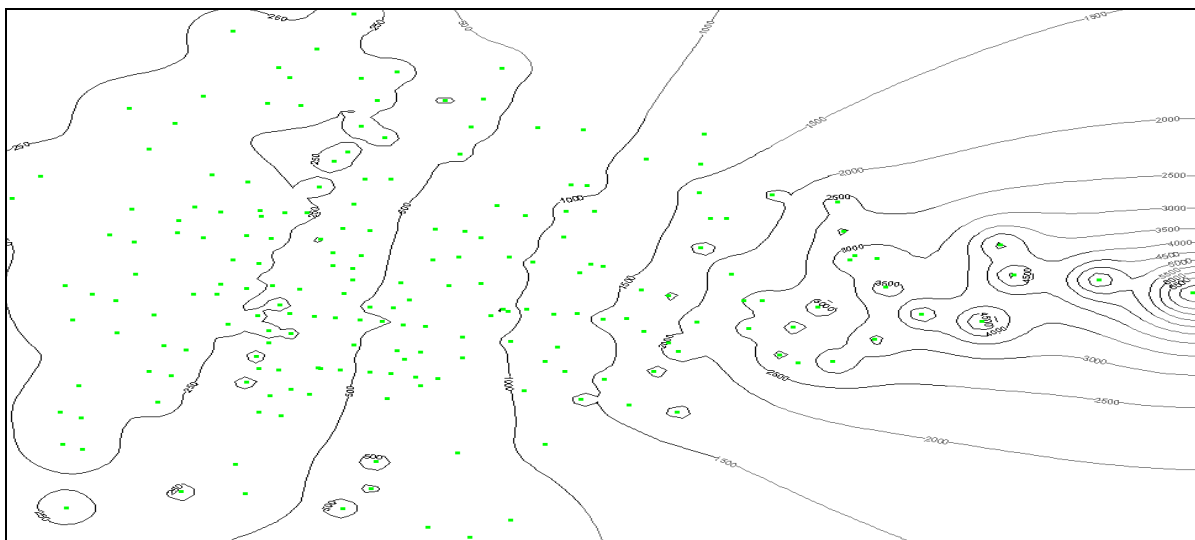


Fig.4.10 – Linhas equi-HiperVolume (12)

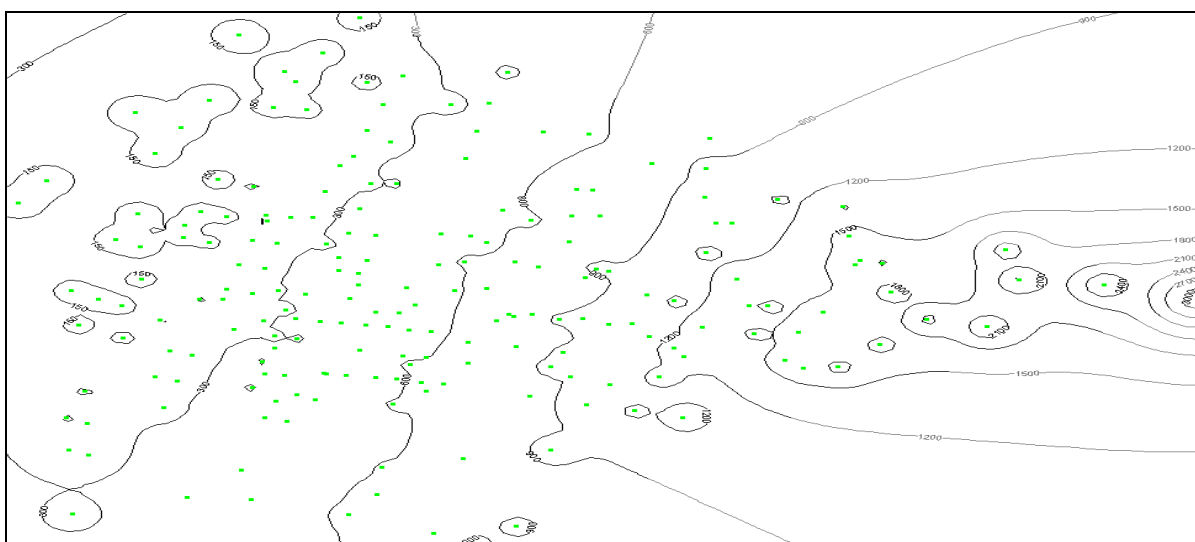


Fig.4.11 – Linhas equi-Soma-Triprodutos (12)

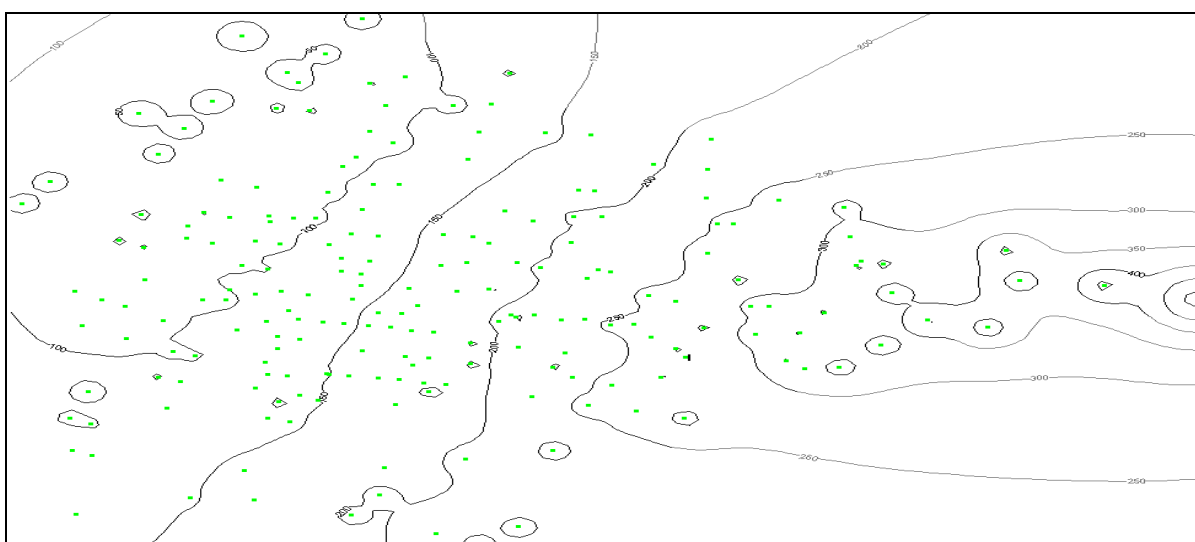


Fig.4.12 – Linhas equi-Soma Biprodutos (12)

As variáveis $X_{23}, X_{21}, X_4, X_8, X_{17}$ e X_3 apresentam as maiores correlações com a segunda componente principal:

	Y2
X1	0,2175
X2	-0,2645
X3	-0,7481
X4	-0,7652
X5	-0,1135
X6	-0,3732
X7	-0,2328
X8	-0,7652
X9	0,2175
X10	-0,5091
X11	-0,3732
X12	-0,2328
X13	0,2628
X14	-0,1554
X15	-0,6017
X16	-0,3773
X17	-0,7581
X18	-0,5698
X19	0,0818
X20	-0,3702
X21	-0,7825
X22	-0,5917
X23	-0,8898
X24	-0,5752
X25	0,5674

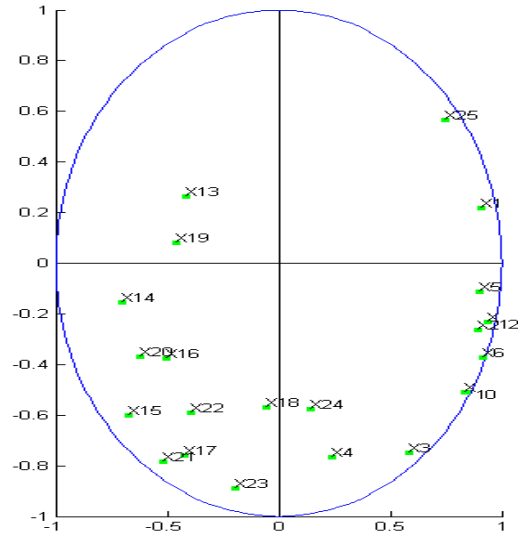


Fig.4.13 – Círculo de Correlações (12)

Opõem-se paralelepípedos com duas dimensões grandes e pelo menos uma das restantes dimensões pequena a paralelepípedos com todas as dimensões pequenas.

Ou seja, X_3 e X_4 são preponderantes porque são, precisamente, estas duas dimensões que distinguem os dois grupos de paralelepípedos: são elevadas num dos grupos e reduzidas no outro.

Do mesmo modo, $X_{23} = X_3 - X_1$ e $X_{17} = \frac{X_3}{X_1}$ diferenciam os dois tipos de paralelepípedos na medida em que para todos eles a primeira dimensão é pequena, mas a terceira dimensão é grande num grupo e pequena no outro. O mesmo para a variável $X_{21} = X_4 - X_1$.

Esta situação é análoga ao caso tridimensional, onde a variável Máximo-Mínimo tinha o maior contributo para a formação da segunda componente.

Nas figuras 4.14 e 4.15 apresentam-se as linhas de nível relativas às variáveis $X_3 - X_1 (= X_{23})$ e X_4 . São, aproximadamente, horizontais e o seu valor aumenta no sentido negativo de segundo eixo.

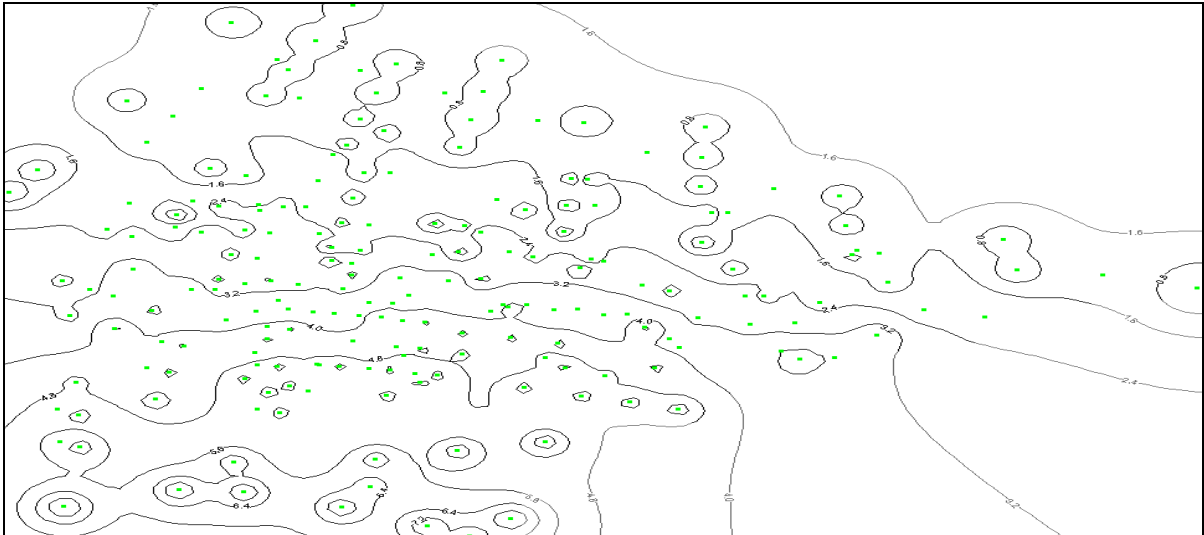


Fig.4.14 – Linhas equi - $X_3 - X_1$ (12)

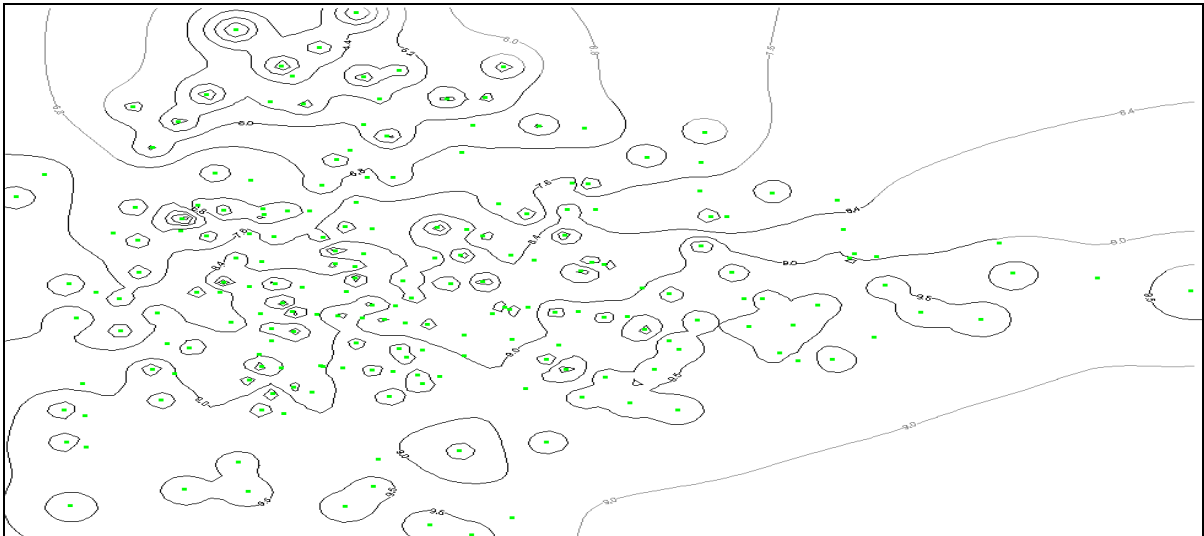


Fig.4.15 – Linhas equi- X_4 (12)

As variáveis que mais contribuem para a formação desta componente não pertencem todas a um mesmo bloco. Na maioria pertencem ao 2º maior bloco, mas existem também variáveis de outros blocos: em particular, $X_4 (= X_8)$ e X_3 .

Para a terceira componente são as variáveis X_{18} e X_{24} que mais contribuem para a sua formação:

	Y3
X1	0,3168
X2	-0,2701
X3	-0,0504
X4	0,5207
X5	0,2256
X6	0,1183
X7	0,1528
X8	0,5207
X9	0,3168
X10	0,1394
X11	0,1183
X12	0,1528
X13	0,4177
X14	0,5142
X15	-0,0776
X16	0,3037
X17	-0,3410
X18	-0,6802
X19	0,6279
X20	0,6636
X21	0,1693
X22	0,2774
X23	-0,3092
X24	-0,6906
X25	0,0287

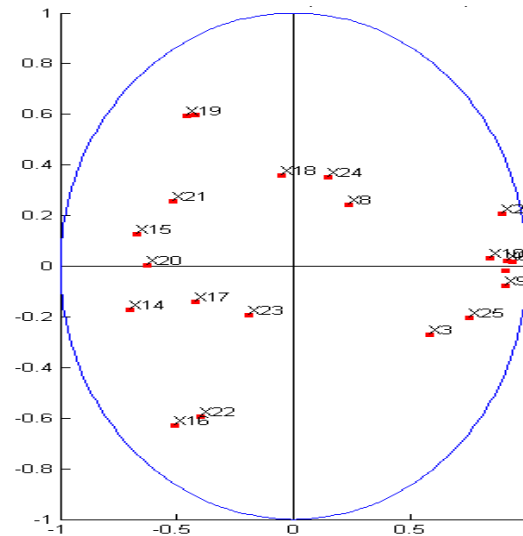


Fig.4.16 -Círculo de Correlações (13)

$X_{18} \left(= \frac{X_2}{X_1} \right)$ e $X_{24} (= X_2 - X_1)$ são as variáveis com maior contributo para a terceira componente.

Porém, não se consegue uma interpretação consequente.

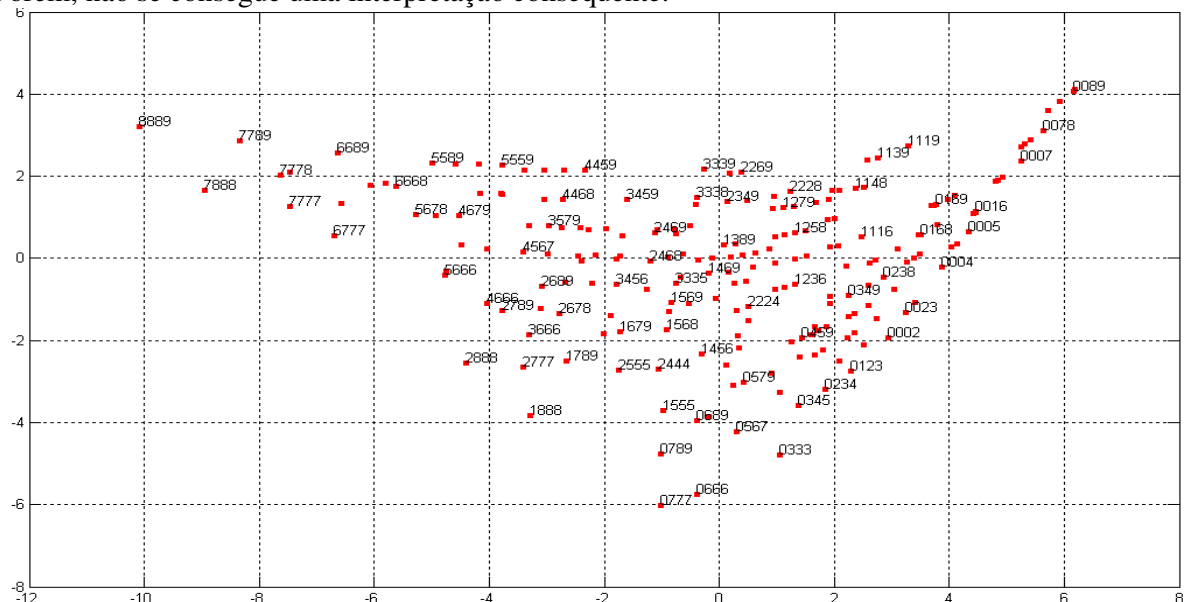


Fig.4.17 – Representação dos indivíduos no plano principal (1,3)

4.3 Caso Contínuo

Foram gerados 220 paralelepípedos rectângulos, cujas dimensões seguem uma lei contínua uniforme]0,10] (Anexo VI). Consideram-se as doze variáveis iniciais.

Para estes paralelepípedos efectuou-se uma análise análoga à que tinha sido efectuada para o primeiro caso discreto.

Tem-se por objectivo testar a qualidade da interpolação de um caso contínuo através do caso discreto.

4.3.1 Estatística Descritiva das Variáveis

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Média	2,4753	5,0096	7,3609	125,56	149,64	7,3609	2,4753	4,9486	24,94	26,9110	4,8856	0,3263
Mediana	2,0535	4,9626	7,7521	69,698	128,385	7,7521	2,0535	5,0279	21,398	3,448	4,8354	0,2900
Moda	#N/D	4,1288	8,7414	#N/D	309,85	8,7414	#N/D	#N/D	#N/D	#N/D	6,0835	#N/D
Variância	3,8847	4,9337	4,2690	21832	11274	4,2690	3,8847	3,0896	313,15	32767	4,5597	0,0548
Desvio-padrão *	1,9710	2,2212	2,0662	147,76	106,18	2,0662	1,9710	1,7577	17,6960	181,0200	2,1353	0,2341
Curtose **	2,8846	2,1456	2,6519	4,8696	3,0656	2,6519	2,8846	2,4243	3,0656	135,8200	2,2327	2,6140
Assimetria	0,8111	0,0806	-0,7240	1,5910	0,8430	-0,7240	0,8111	-0,0645	0,8430	11,2140	0,0311	0,6398
Mínimo	0,0039	0,4417	1,6459	0,0461	3,7788	1,6459	0,0039	1,0505	0,6298	1,0301	0,2303	0,0004
Máximo	7,9781	9,8376	9,9746	624,56	447,74	9,9746	7,9781	8,7135	74,6230	2358	9,5195	0,9708
Coefficiente de variação	0,7963	0,4434	0,2807	1,1768	0,7096	0,2807	0,7963	0,3552	0,7095	6,7266	0,4371	0,7174

* - Desvio Padrão Corrigido

** - Definição usada: $k = \frac{E(x-\mu)^4}{\sigma^4}$

Tabela 4.8

Comparando com o caso discreto, verifica-se que a única variável onde as diferenças são significativas é a variável $X_{10} = \frac{\text{Máximo}}{\text{Mínimo}}$. Esta diferença deve-se ao facto de as dimensões de estes paralelepípedos poderem assumir uma das dimensões muito próximas de zero, o que pode tornar os valores desta variável extremamente elevados.

Tal como no caso discreto, também se efectuam os cálculos para a média e variância das variáveis X_1 e X_3 . Neste caso, considera-se que $X_1, X_2, X_3 \sim U_{[0,10]}$.

$$\text{Portanto, } F_{X_1}(x) = 1 - \left(1 - \frac{x}{10}\right)^3 \text{ e } F_{X_3}(x) = \left(\frac{x}{10}\right)^3$$

Logo,

$$E(X_1) = \int_{-\infty}^{+\infty} x f_{X_1}(x) dx = \int_0^{10} x \frac{3}{10} \left(1 - \frac{x}{10}\right)^2 dx = 2,5$$

$$V(X_1) = E(X_1^2) - E^2(X_1) = \int_{-\infty}^{+\infty} x^2 f_{X_1}(x) dx - 2,5^2 = \int_1^{10} x^2 \frac{3}{10} \left(1 - \frac{x}{10}\right)^2 dx - 2,5^2 = 3,75$$

$$E(X_3) = \int_{-\infty}^{+\infty} x f_{X_3}(x) dx = \int_0^{10} 3x \frac{x^2}{10^3} dx = 7,5$$

$$V(X_3) = E(X_3^2) - E^2(X_3) = \int_{-\infty}^{+\infty} x^2 f_{X_3}(x) dx - 7,5^2 = \int_0^{10} 3x^2 \frac{x^2}{10^3} dx - 7,5^2 = 3,75$$

4.3.2 Análise em Componentes Principais

4.3.2.1 Matriz de Correlações

Analisando a matriz de correlações obtida (Tabela 4.9) conclui-se não existirem diferenças expressivas em relação ao caso discreto, à exceção da variável X_{10} cujas alterações nas correlações com as restantes variáveis são bastante significativas, pelo motivo atrás referido.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1	0,58	0,44	0,92	0,88	0,44	1	0,79	0,88	-0,17	-0,50	0,92
X2	0,58	1	0,66	0,73	0,86	0,66	0,58	0,90	0,86	0,05	0,10	0,41
X3	0,44	0,66	1	0,52	0,69	1	0,44	0,83	0,69	0,02	0,56	0,15
X4	0,92	0,73	0,52	1	0,96	0,52	0,92	0,85	0,96	-0,11	-0,34	0,76
X5	0,88	0,86	0,69	0,96	1,0	0,69	0,88	0,96	1,00	-0,06	-0,14	0,69
X6	0,44	0,66	1	0,52	0,69	1	0,44	0,83	0,69	0,02	0,56	0,15
X7	1	0,58	0,44	0,92	0,88	0,44	1	0,79	0,88	-0,17	-0,50	0,92
X8	0,79	0,90	0,83	0,85	0,96	0,83	0,79	1	0,96	-0,03	0,08	0,58
X9	0,88	0,86	0,69	0,96	1	0,69	0,88	0,96	1	-0,06	-0,14	0,69
X10	-0,17	0,05	0,02	-0,11	-0,06	0,02	-0,17	-0,03	-0,06	1	0,17	-0,18
X11	-0,50	0,10	0,56	-0,34	-0,14	0,56	-0,50	0,08	-0,14	0,17	1	-0,71
X12	0,92	0,41	0,15	0,76	0,69	0,15	0,92	0,58	0,69	-0,18	-0,71	1

Tabela 4.9

Continuam a poder distinguir-se os mesmos três blocos de variáveis fortemente correlacionadas: analíticas, função dos extremos e máximo.

	X5	X9	X8	X4	X1	X7	X2	X12	X10	X11	X3	X6
X5	1	1	0,96	0,96	0,88	0,88	0,86	0,69	-0,06	-0,14	0,69	0,69
X9	1	1	0,96	0,96	0,88	0,88	0,86	0,69	-0,06	-0,14	0,69	0,69
X8	0,96	0,96	1	0,85	0,79	0,79	0,90	0,58	-0,03	0,08	0,83	0,83
X4	0,96	0,96	0,85	1	0,92	0,92	0,73	0,76	-0,11	-0,34	0,52	0,52
X1	0,88	0,88	0,79	0,92	1	1	0,58	0,92	-0,17	-0,50	0,44	0,44
X7	0,88	0,88	0,79	0,92	1	1	0,58	0,92	-0,17	-0,50	0,44	0,44
X2	0,86	0,86	0,90	0,73	0,58	0,58	1	0,408	0,05	0,10	0,66	0,66
X12	0,69	0,69	0,58	0,76	0,92	0,92	0,41	1	-0,18	-0,71	0,15	0,15
X10	-0,06	-0,06	-0,03	-0,11	-0,17	-0,17	0,05	-0,18	1	0,17	0,02	0,02
X11	-0,14	-0,14	0,08	-0,34	-0,50	-0,50	0,10	-0,71	0,17	1	0,56	0,56
X3	0,69	0,69	0,83	0,52	0,44	0,44	0,66	0,15	0,02	0,56	1	1
X6	0,69	0,69	0,83	0,52	0,44	0,44	0,66	0,15	0,02	0,56	1	1

Tabela 4.10

4.3.2.2 Valores Próprios da Matriz de Correlações

Na tabela seguinte estão apresentados os valores próprios obtidos para a matriz de correlações:

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
7,7270	64,39	64,39
2,7116	22,60	86,99
0,9684	8,07	95,06
0,4215	3,51	98,57
0,1495	1,25	99,82
0,0186	0,16	99,98
0,0035	0,03	100
0,0000	0,00	100
0,0000	0,00	100
0,0000	0,00	100
0,0000	0,00	100
0,0000	0,00	100

Tabela 4.11

Não existem alterações significativas em relação ao exemplo discreto. O segundo eixo perde algum peso para o terceiro eixo, embora, em conjunto, os dois primeiros eixos expliquem 87% da inércia total (contra os 93% do caso discreto).

4.3.2.3 Interpretação das Componentes Principais

Apresenta-se abaixo a representação gráfica dos indivíduos no primeira plano principal:

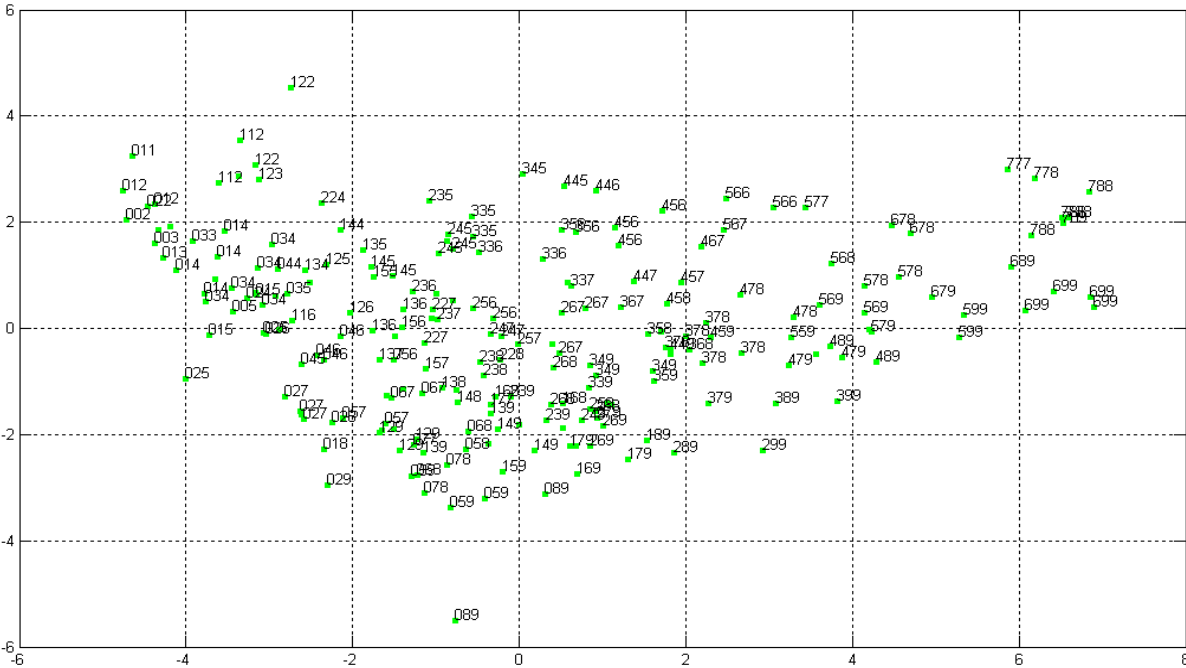


Fig.4.18-Indivíduos representados no plano principal (1,2)

Cada paralelepípedo de dimensões onde a , b e c é representado por $[a][b][c]$, ou seja, pelo maior inteiro que não excede a dimensão.

4.3.2.3.1 Interpretação da Primeira Componente Principal

Observando a figura 4.18 observa-se que no primeiro eixo se encontram em oposição paralelepípedos de grandes dimensões (à direita) e paralelepípedos de pequenas dimensões (à esquerda). Daqui resulta que exista no primeiro eixo uma oposição entre paralelepípedos com valores elevados e reduzidos de área e volume.

As representações das linhas equi-volume e equi-área são ilustrativas desta situação:

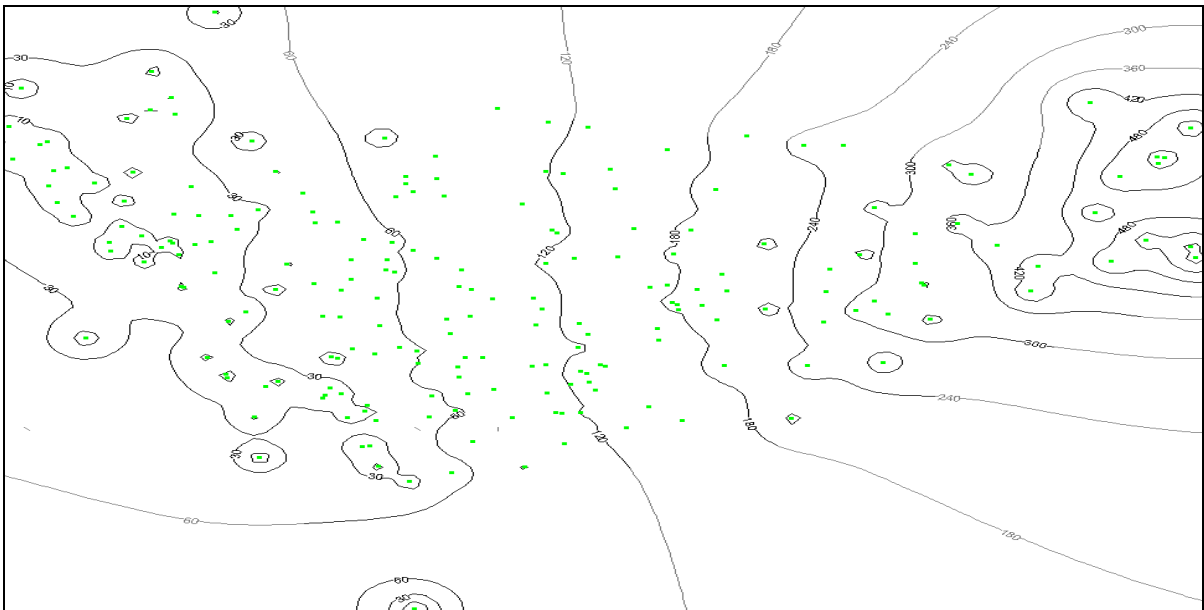


Fig.4.19 – Linhas equi-volume (12)

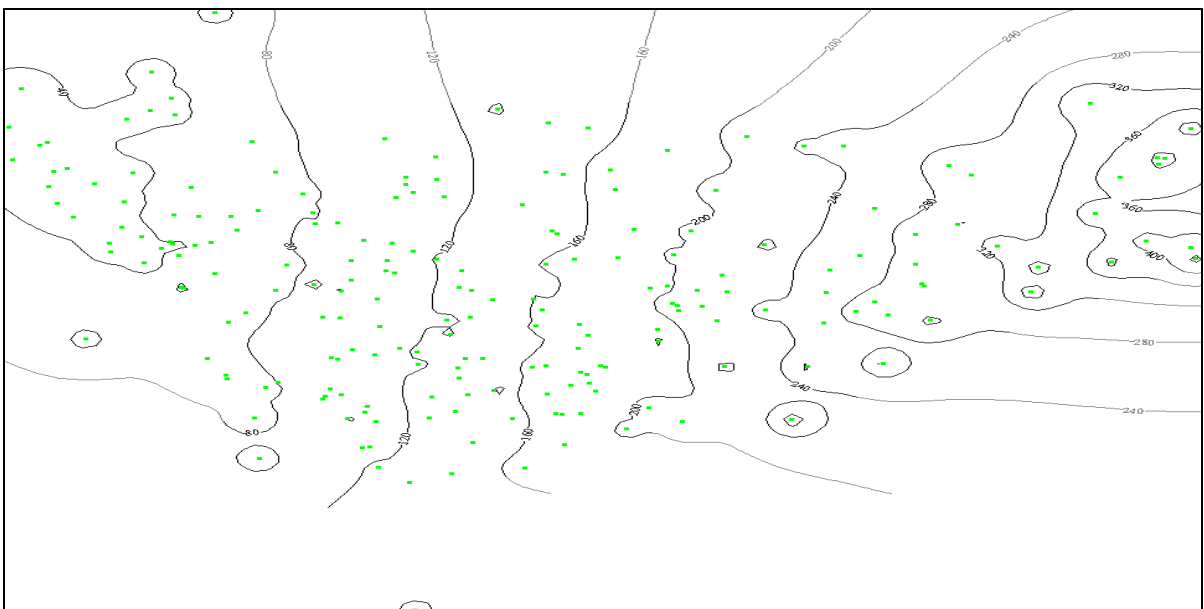


Fig.4.20 – Linhas equi-área (12)

Em ambas as figuras, as linhas vão em crescendo no sentido positivo do primeiro eixo. Na parte esquerda das figuras encontram-se vales mais ou menos profundos evoluindo para picos escarpados. Note-se, no entanto, que nas linhas de nível mais à esquerda não se encontra a verticalidade do caso discreto talvez por haver uma maior possibilidade de valores para as dimensões.

O comportamento dos indivíduos no primeiro eixo no caso contínuo é, assim, análogo ao dos indivíduos no caso discreto, embora neste último caso se verificasse uma '*maior organização*' dos indivíduos. Ainda que mais atenuada, continua patente a estrutura triangular.

Sobrepondo o triângulo obtido no exemplo discreto, verifica-se que todas as arestas são menos marcadas, embora a sua organização e interpretação seja semelhante.

Neste caso, não existem verdadeiros cubos, apenas quase cubos. Esta situação prende-se, evidentemente, com a forma de geração destes paralelepípedos.

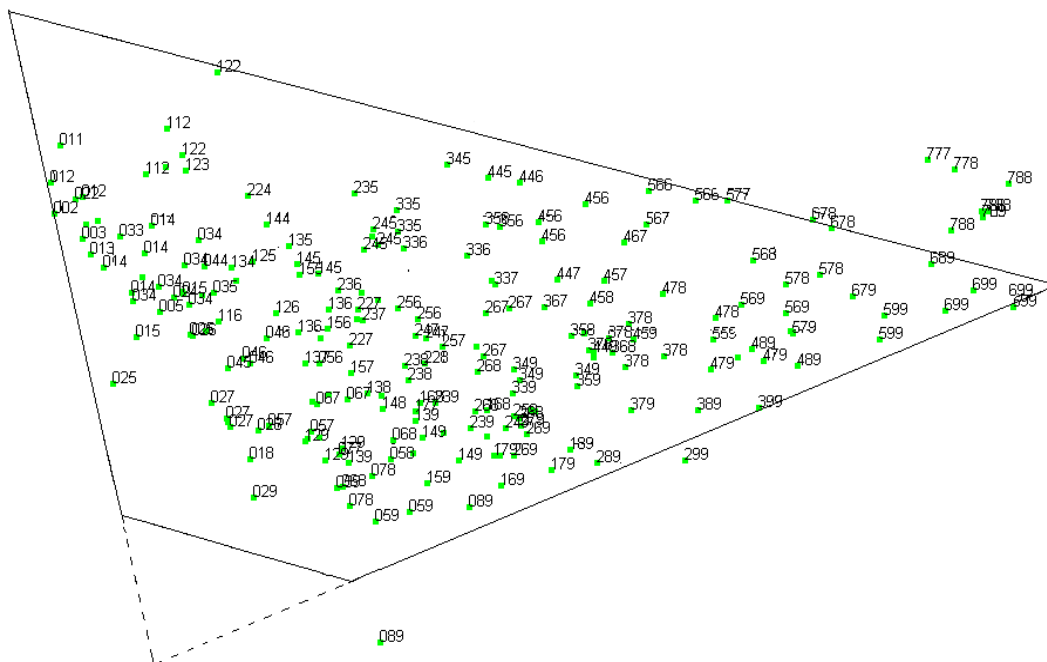


Fig.4.21 – Organização dos Indivíduos

Quer as diferenças na organização dos indivíduos, quer as diferenças nas curvas de nível são pouco significativas, pelo que o *caso discreto parece ser uma boa forma de aproximação do caso contínuo*.

Observando as correlações da primeira componente com as variáveis consideradas verifica-se que as que mais contribuem para a sua formação são as variáveis *volume*, *área total* e *dimensão mínima*, tal como no caso discreto.

Inclusive, a organização do círculo de correlações (Fig.4.22) não apresenta alterações expressivas. X_{10} apresenta-se mais no interior do círculo (devido ao mínimo ser próximo de zero) e também X_3 e X_{12} se aproximam das variáveis de tamanho.

	Y1
X1	0,9172
X2	0,8225
X3	0,7039
X4	0,9470
X5	0,9890
X6	0,7039
X7	0,9172
X8	0,9651
X9	0,9890
X10	-0,0991
X11	-0,1655
X12	0,7453

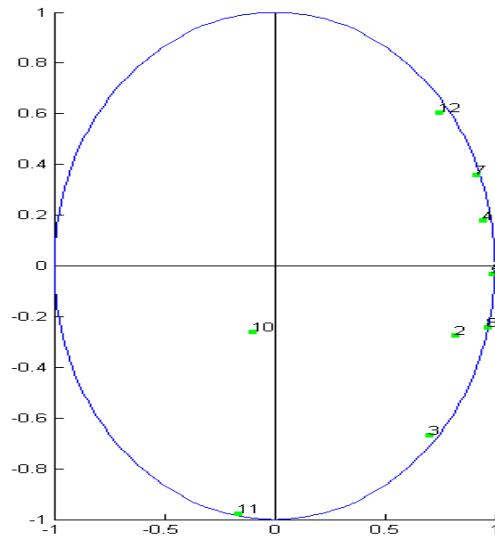


Fig.4.22 – Círculo de Correlações (12)

4.3.2.3.2 Interpretação da Segunda Componente Principal

No segundo eixo principal encontram-se em oposição, tal como no caso discreto, paralelepípedos que tem pelo menos duas componentes iguais ou muito próximas com os têm pelo menos duas muito diferentes.

A análise das correlações mostra que a variável *Máximo – Mínimo* é dominante para a formação desta componente, bem como a variável dimensão máxima, embora com menor relevância.

	Y2
X1	0,3577
X2	-0,2740
X3	-0,6696
X4	0,1786
X5	-0,0305
X6	-0,6696
X7	0,3577
X8	-0,2441
X9	-0,0305
X10	-0,2607
X11	-0,9781
X12	0,6050

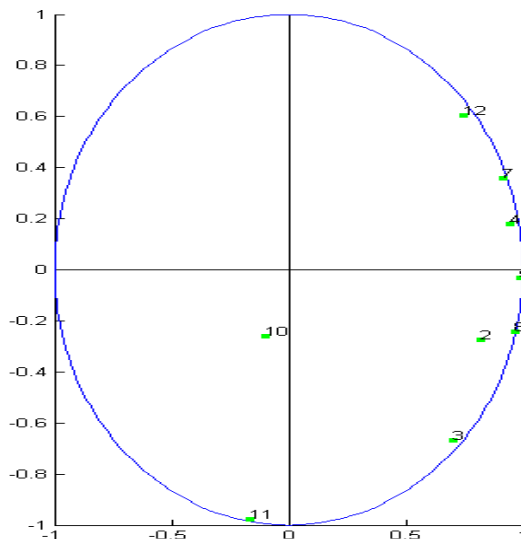


Fig.4.23 - Círculo de Correlações (12)

No caso discreto, distinguíam-se, relativamente à segunda componente três grupos de variáveis: $\{X_3, X_6, X_{10}, X_{11}\}, \{X_{12}\}$ e $\{X_1, X_2, X_4, X_5, X_7, X_8, X_9\}$. Neste caso passam a distinguir-se três novos grupos de variáveis: $\{X_{11}\}, \{X_3, X_6, X_{12}\}$ e $\{X_1, X_2, X_4, X_5, X_7, X_8, X_9, X_{10}\}$, caracterizados por correlações fortes, intermédias e fracas com a segunda componente, respectivamente. Destaca-se o deslocamento de X_{10} para o interior do círculo e também a diminuição das correlações de $X_3 (= X_6)$.

Observando as linhas equi-Máximo-Mínimo verifica-se que o valor que estas representam aumenta no sentido negativo do 2º eixo, onde se encontram os paralelepípedos duas dimensões muito diferentes:

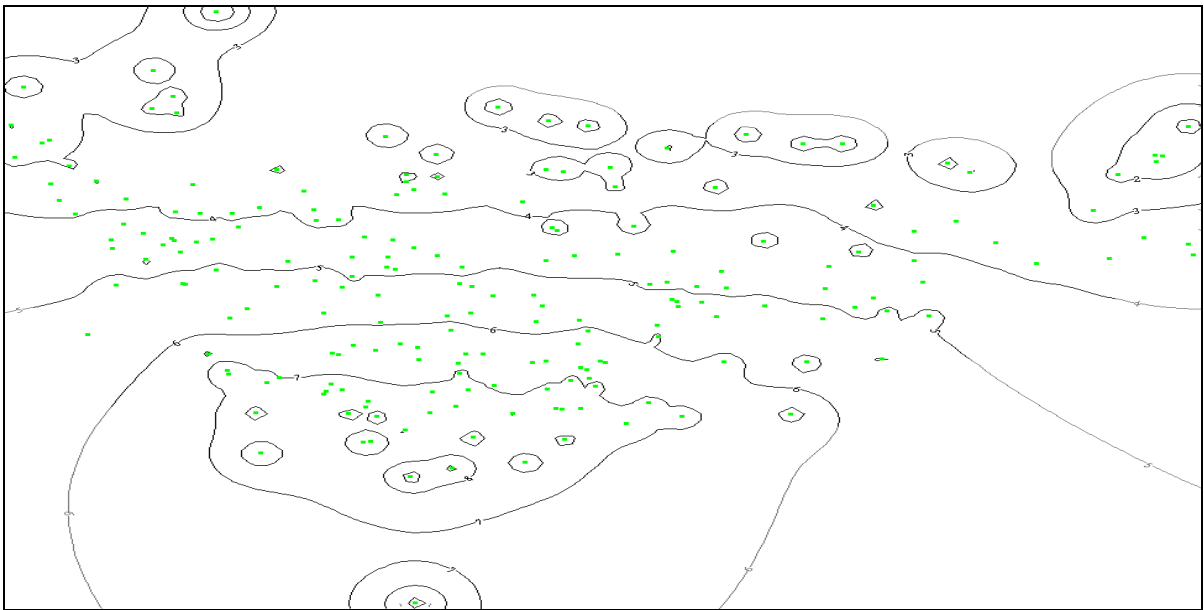


Fig.4.24 – Linhas equi- Máximo-Mínimo (12)

O teste das curvas de nível permite-nos novamente inferir que o caso discreto representa muito bem o caso contínuo.

4.3.2.3.3 Interpretação da Terceira Componente Principal

No caso discreto, o terceiro eixo permitia, a partir da segunda dimensão, fazer a distinção entre ‘palitos’ e ‘tábuas’. Agora, tal distinção é feita pelo 4º componente, que explica apenas 3,5% da inércia total.

A variável que mais contribui para a formação do terceira componente é X_{10} , que possui com esta uma correlação de 95%, sendo que as restantes têm uma correlação praticamente nula (Fig.4.25).

	Y3
X1	-0,0039
X2	0,1269
X3	-0,1178
X4	0,0350
X5	0,0425
X6	-0,1178
X7	-0,0039
X8	0,0058
X9	0,0425
X10	0,9521
X11	-0,1105
X12	0,0292

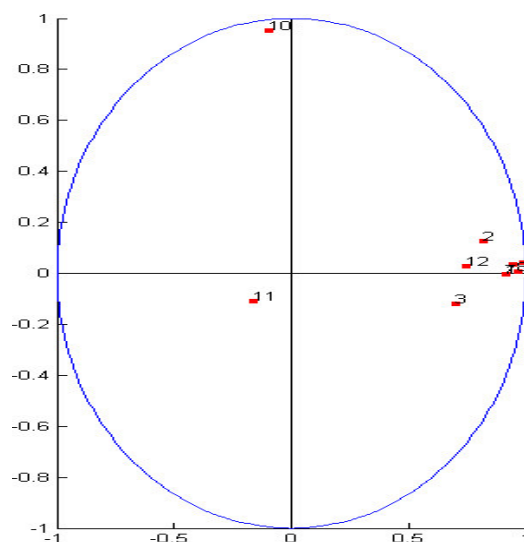


Tabela 4.25 – Círculo de Correlações (13)

Existem dois paralelepípedos que se destacam evidentemente na formação deste eixo. São paralelepípedos que se caracterizam por ter uma dimensão elevada e cuja dimensão mínima é muito reduzida o que provoca um valor alto na variável X_{10} .

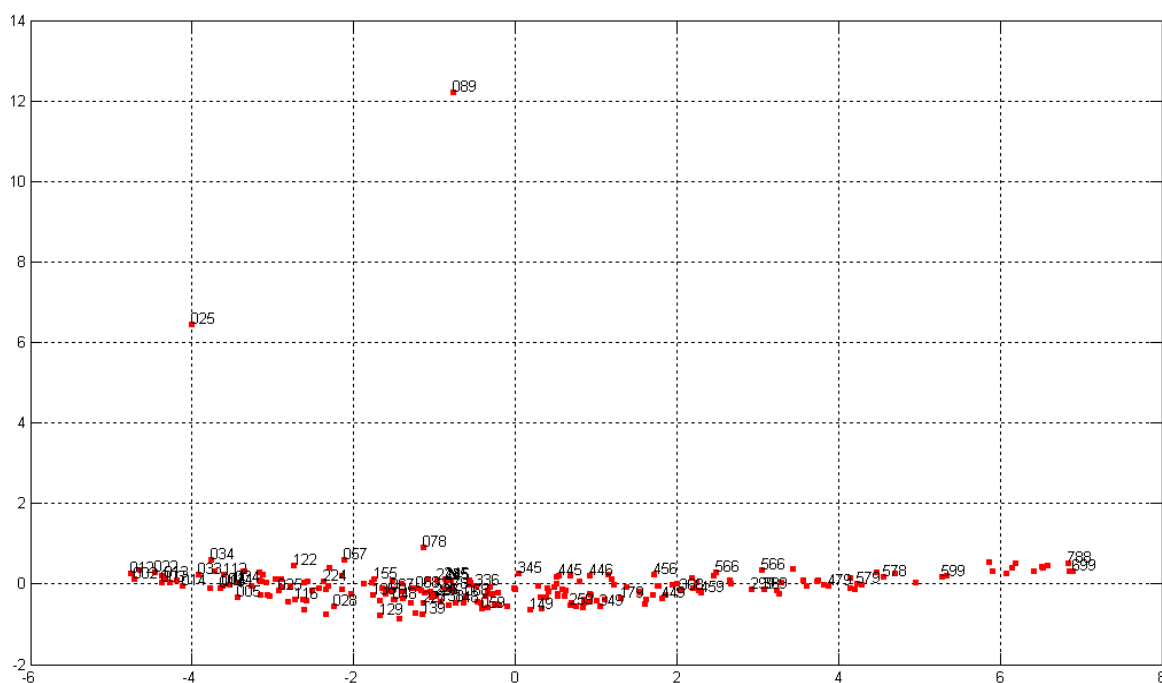


Fig.4.26 – Indivíduos representados no plano principal (1,3)

Realizando ACP excluindo os dois anteriores paralelepípedos resulta a seguinte representação.

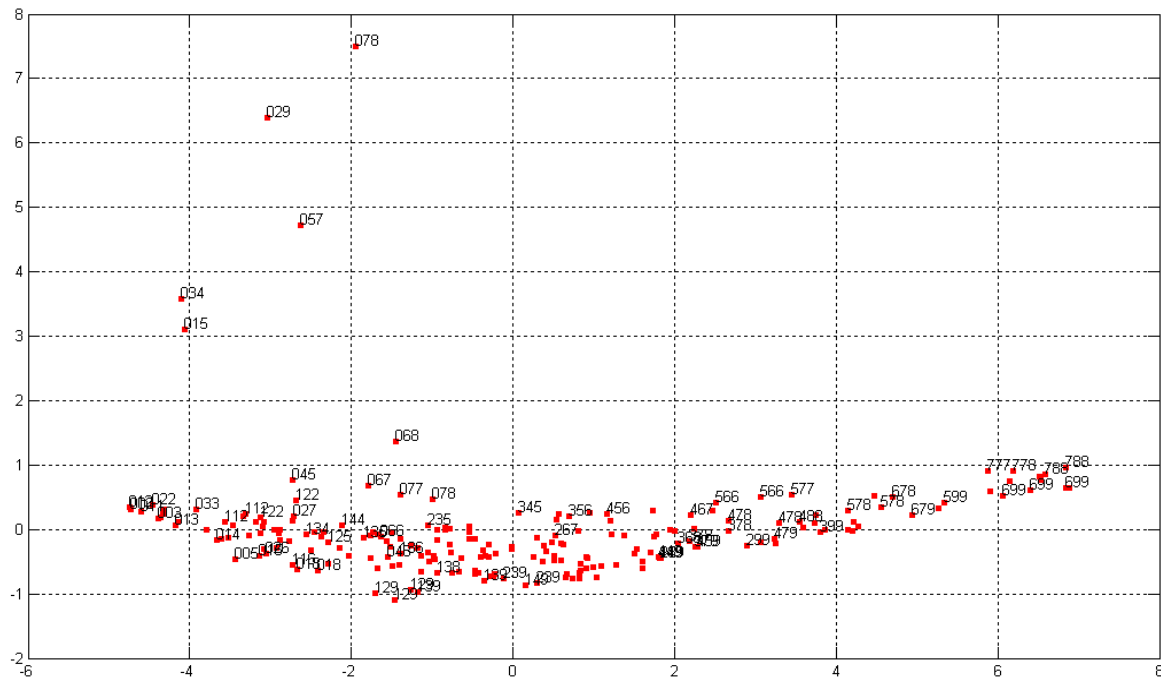


Fig.4.27 – Indivíduos representados no plano principal (1,3)

Este tipo de efeito já havia sido encontrado no estudo de perfis de consumidores (Anexo I, página 124).

Assim, o terceiro eixo coloca em evidência paralelepípedos em que uma dimensão é muito elevada e a dimensão mínima muito reduzida.

Embora a terceira componente tenha nova interpretação, fruto dos valores obtidos para a dimensão mínima, nas duas primeiras componentes, que são as que mais contribuem para a inércia explicada, as diferenças discreto, contínuo não são significativas.

4.3.3 Estabilidade de ACP realizada

4.3.3.1 Variação do primeiro valor próprio em função da dimensão

Para o estudo da interferência da dimensão da amostra no primeiro valor próprio efectuou-se o mesmo estudo do caso discreto.

Recolheram-se, ao acaso, 30 amostras de dimensões sucessivamente 10, 20, 30, ... 210, onde cada amostra foi obtida sem reposição.

Observa-se uma muito rápida aproximação do valor assintótico ($\lambda = 7,727$) com o aumentar da dimensão da amostra (Fig.4.28).

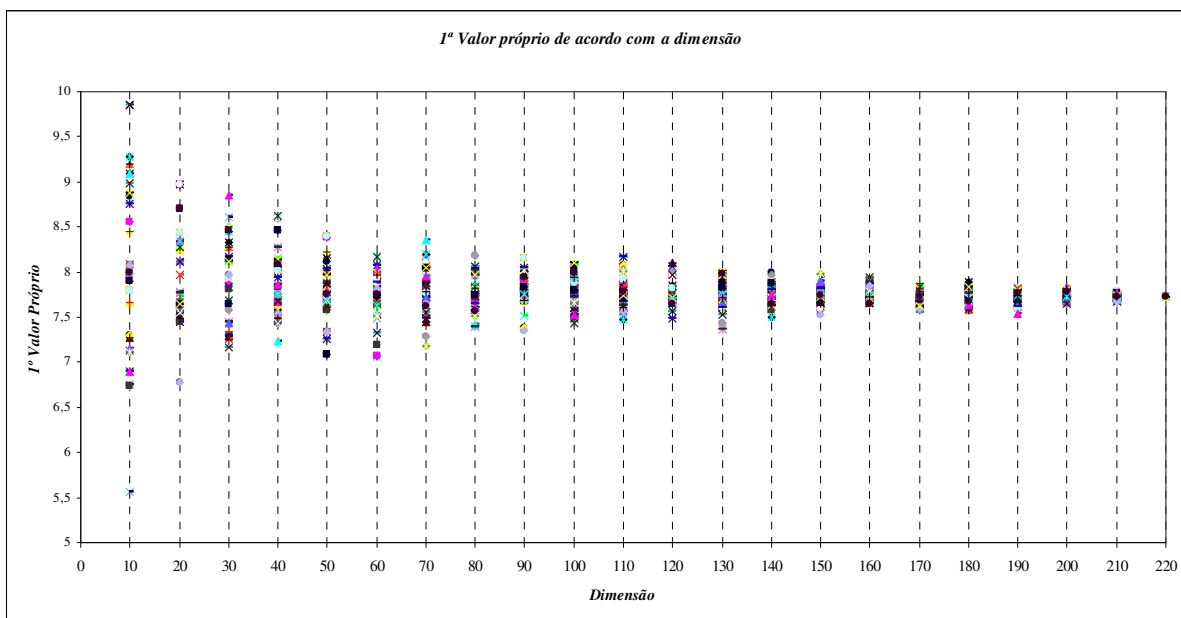


Fig.4.28 -Variação do primeiro valor próprio em função da dimensão

Verifica-se também uma maior variabilidade do primeiro valor próprio para amostras pequenas, que vai estabilizando para amostras maiores.

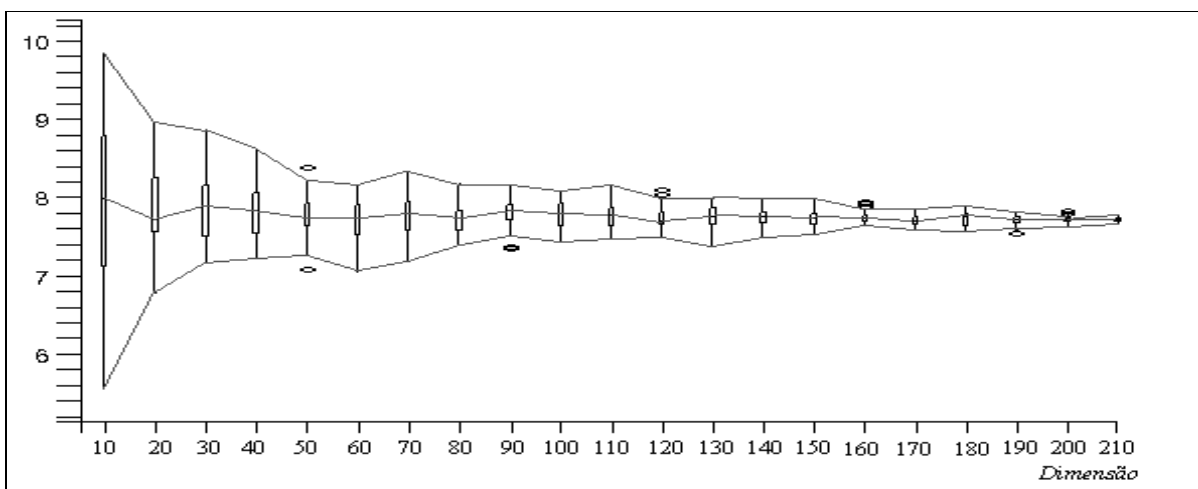


Fig.4.29-Diagrama caixa de bigodes

À excepção da dimensão 10, onde o coeficiente de variação é superior ao caso discreto (12%), nas dimensões seguintes o coeficiente de correlação é ligeiramente inferior (Tabela 4.12).

Na figura 4.30 comparam-se os valores médios do primeiro valor próprio no caso discreto e caso contínuo.

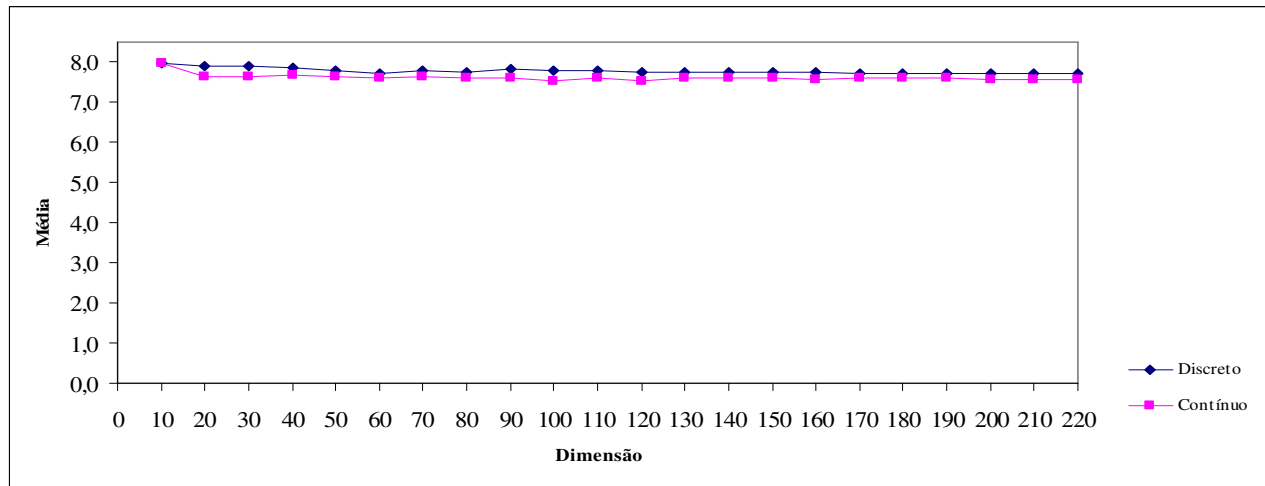


Fig.4.30 – Comparação do caso discreto e caso contínuo (valores médios)

Claramente, as duas situações são idênticas.

Análise da normalidade do primeiro valor próprio

A análise da normalidade do primeiro valor próprio foi feita, como anteriormente, aplicando o teste (γ_1, γ_2) .

Das 30 amostras obtiveram-se as seguintes estatísticas:

Dimensão	Média	Desvio Padrão	Assimetria	Curtose	Coefficiente de Variação
10	7,975	0,973	-0,240	2,659	0,1220
20	7,895	0,492	0,462	3,092	0,0623
30	7,911	0,445	0,115	2,167	0,0563
40	7,859	0,350	0,488	2,657	0,0446
50	7,780	0,305	-0,005	3,084	0,0391
60	7,709	0,283	-0,734	3,022	0,0367
70	7,782	0,267	-0,182	2,759	0,0343
80	7,751	0,194	0,239	2,395	0,0250
90	7,815	0,176	-0,857	3,985	0,0225
100	7,789	0,186	-0,213	2,022	0,0239
110	7,781	0,161	0,466	2,968	0,0207
120	7,757	0,143	0,860	3,305	0,0184
130	7,762	0,160	-0,499	2,938	0,0206
140	7,755	0,112	-0,037	3,077	0,0144
150	7,750	0,099	0,166	3,095	0,0128
160	7,757	0,080	0,813	3,121	0,0103
170	7,710	0,072	0,205	2,407	0,0094
180	7,733	0,081	-0,183	2,051	0,0105
190	7,719	0,069	-0,653	3,101	0,0089
200	7,729	0,046	0,227	2,735	0,0059
210	7,722	0,029	0,112	2,309	0,0038

Tabela 4.12

A representação Assimetria vs Curtose (Fig.4.31) para as diferentes dimensões mostra que a distribuição amostral do 1º valor próprio passa testes de normalidade, com exceção de $n = 90$, $n = 120$ e $n = 190$. Para $n = 60$ e $n = 160$ o teste é indecidível.

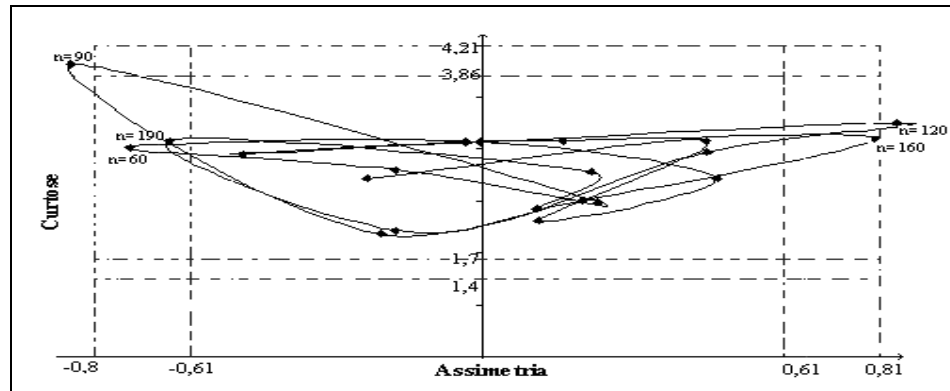


Fig.4.31 – Representação Assimetria vs Curtose

Apenas 3 valores em 21 (14%) não passam o teste de normalidade, o que significa uma ligeira melhoria em relação ao caso discreto (25 %)

As conclusões são, todavia, análogas.

Em relação, à variação do primeiro valor em função da dimensão da amostra quando se retira uma variável ou grupos de variáveis, as conclusões são as mesmas do caso discreto.

O valor próprio altera-se consoante a variável ou grupo de variáveis retirado, mas a aproximação ao valor assintótico ocorre do mesmo modo que considerando a totalidade das variáveis

Apresenta-se o que sucede retirando a variável X_4 e compara-se com a variação tomando a totalidade das variáveis e com a variação retirando a mesma variável no caso discreto.

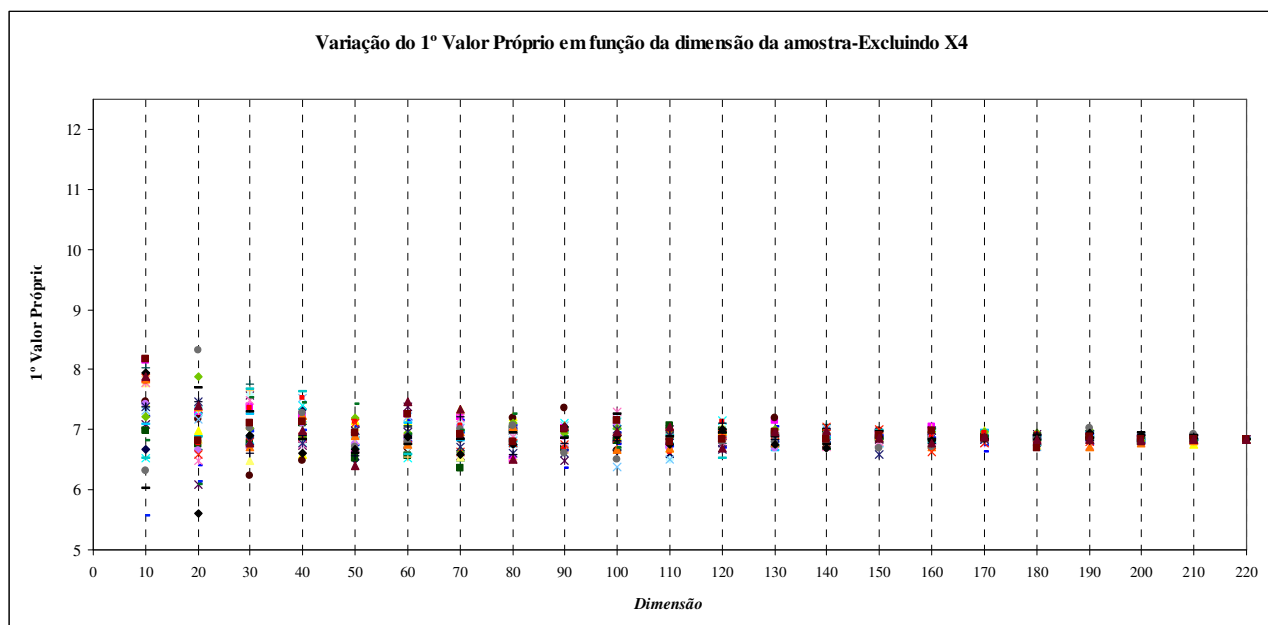


Fig.4.32 – Variação do Primeiro Valor Próprio em função da dimensão da amostra

Se se confrontarem (Fig.4.33) os valores médios para cada dimensão, com todas as variáveis e omitindo X_4 , observa-se que de facto o comportamento na aproximação do valor próprio real, para cada situação, não se altera significativamente.

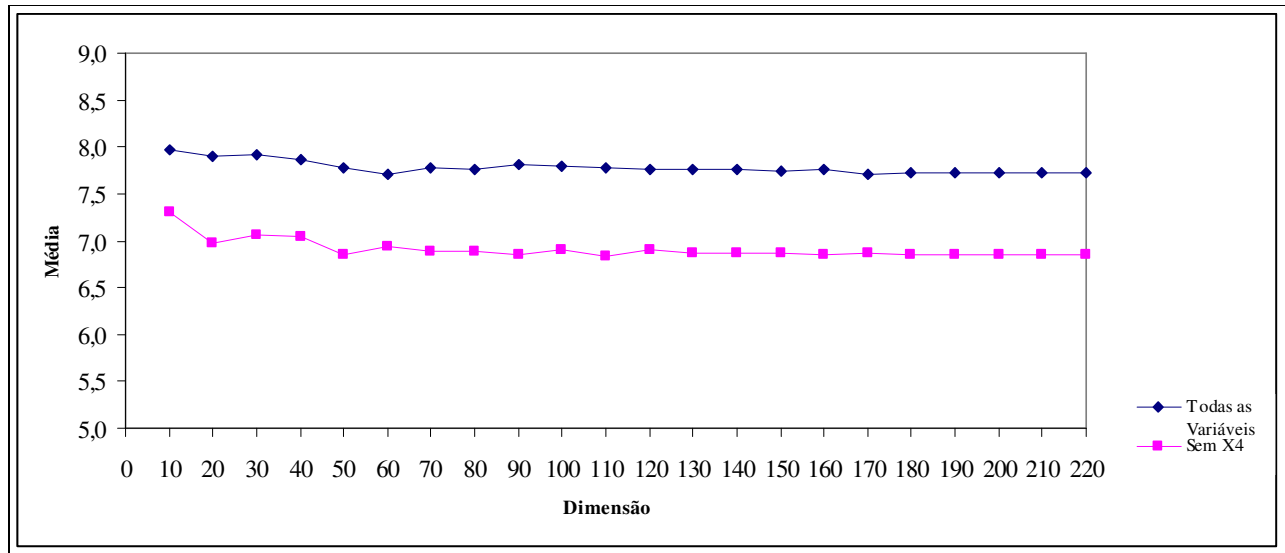


Fig.4.33 – Comparação de Valores Médios do Primeiro Valor Próprio

Tomando os valores médios do primeiro valor próprio para cada dimensão, retirando X_4 quer no caso discreto, quer no caso contínuo, tem-se que as duas situações praticamente se sobrepõem (Fig.4.34)

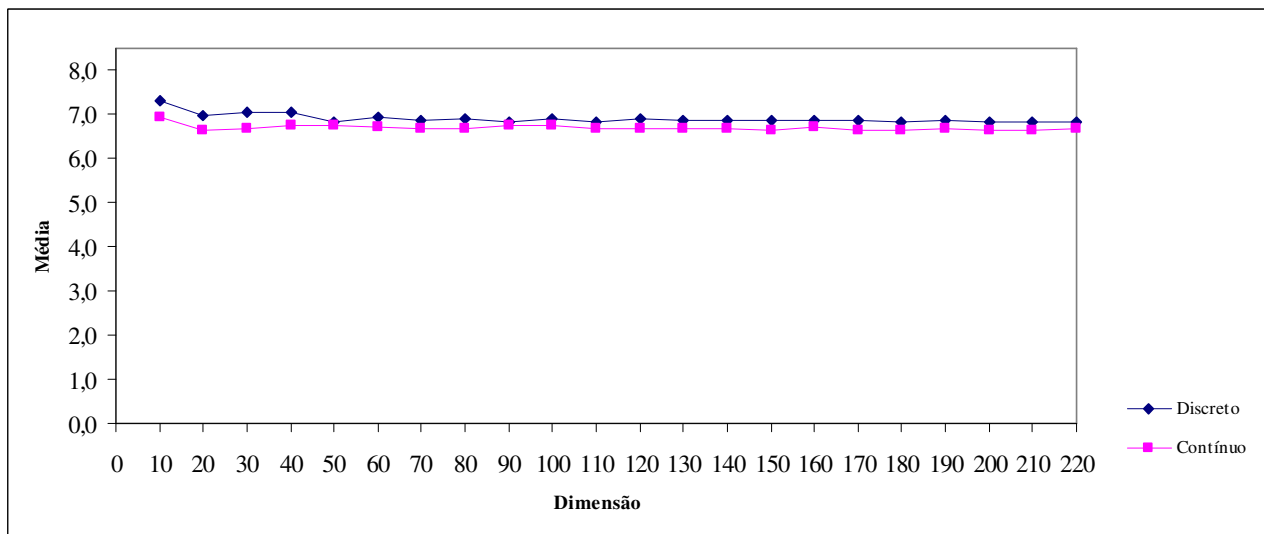


Fig.4.34 – Comparação dos Valores Médios do Primeiro Valor Próprio no caso discreto e contínuo

4.3.3.2 Variações Angulares

O gráfico seguinte refere-se à variação do ângulo entre o primeiro eixo real e o ‘primeiro eixo amostral’ em função da dimensão.

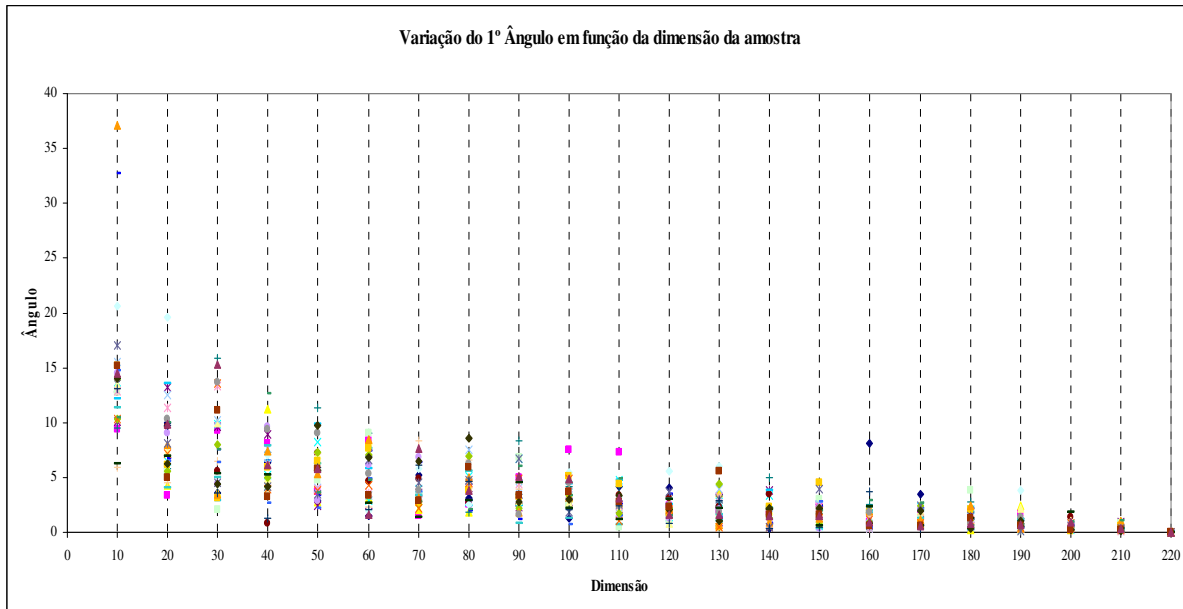


Fig.4.35 – Variação do 1º ângulo em função da dimensão da amostra

Embora para amostras pequenas a variação do ângulo seja muito grande (desde um mínimo $6,2177^\circ$ até um de máximo de $37,094^\circ$), à semelhança do caso discreto, a estabilidade do sistema aumenta com o aumentar da dimensão das amostras, uma vez que o ângulo vai decrescendo monotonicamente (em termos médios). Para amostras maiores ou iguais a 60, tem-se que o ângulo máximo obtido é sempre inferior a 10° (em particular, inferior a $9,1^\circ$).

Neste caso, verifica-se que o ângulo nunca ultrapassa o limite de estabilidade de 45° .

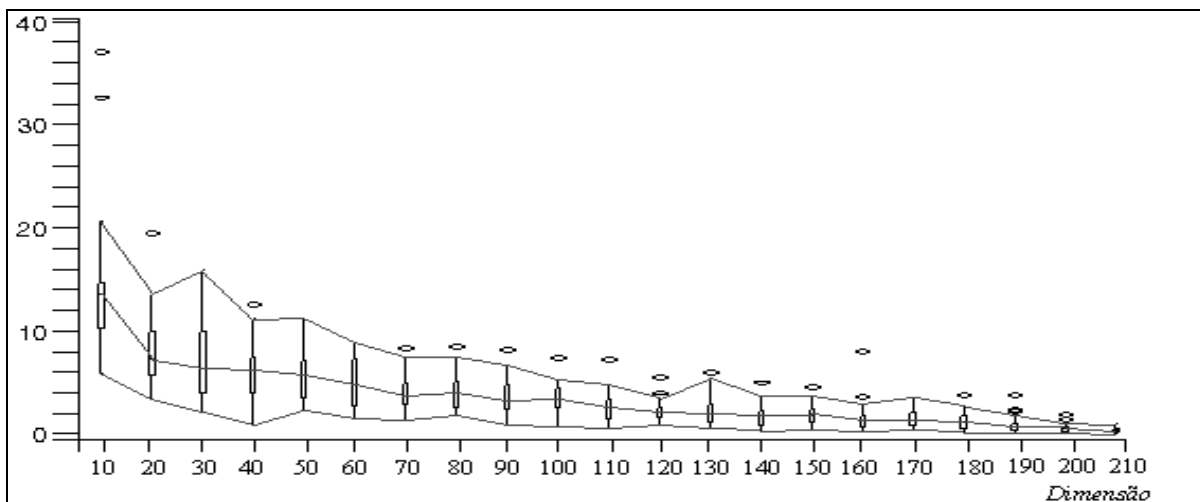


Fig.4.36 – Diagrama caixa de bigodes

Quanto à variação do ângulo entre o segundo eixo real e o ‘segundo eixo amostral’ em função da dimensão:

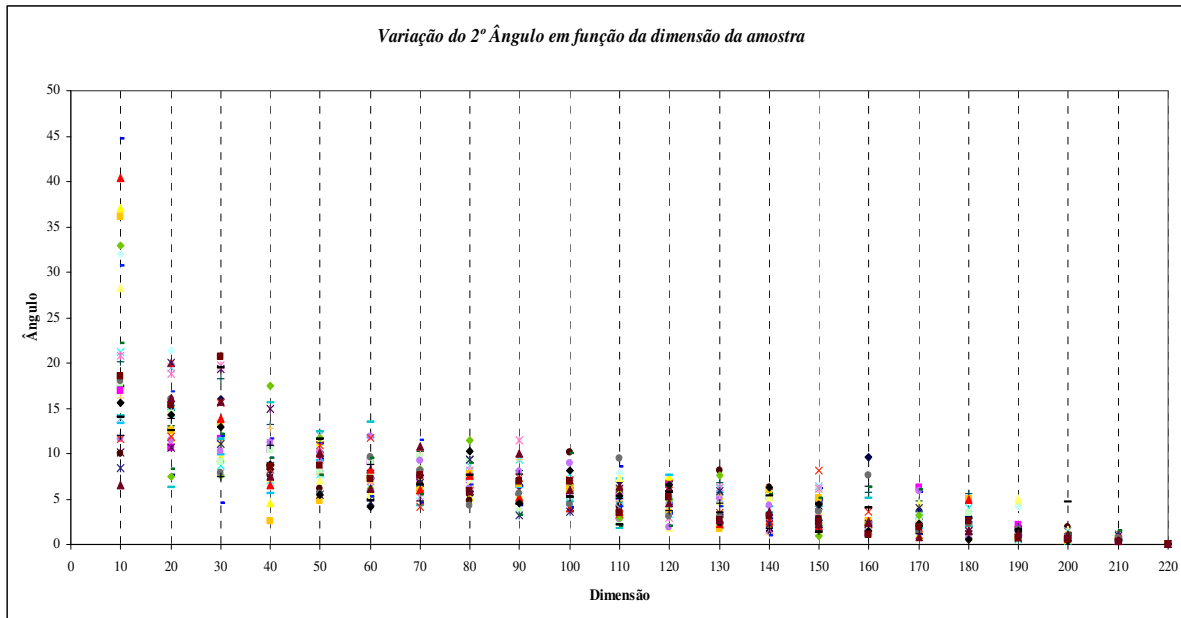


Fig.4.37 – Variação do 2º ângulo em função da dimensão da amostra

Embora estabilize mais lentamente (apenas para $n = 110$ o ângulo máximo obtido é sempre inferior a 10°), o 2º ângulo também vai, naturalmente, diminuindo com o aumentar da amostra. Unicamente para amostras de dimensão 10, o ângulo se aproxima do limite de estabilidade: $44,725^\circ$ (e uma única vez).

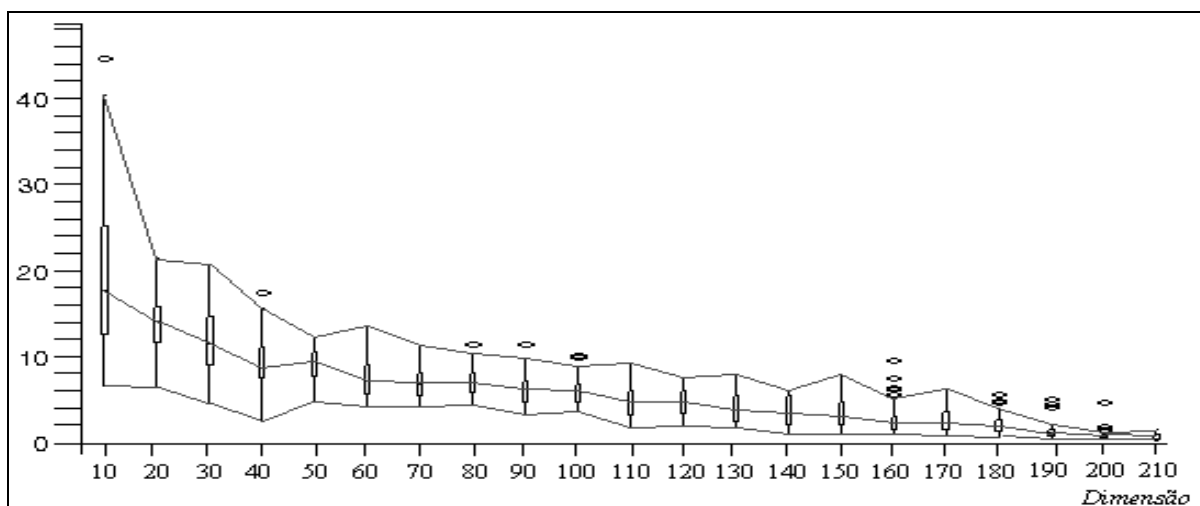


Fig.4.38 – Diagrama caixa de bigodes

Para a variação do ângulo entre o terceiro eixo real e o ‘terceiro eixo amostral’ verifica-se que para pequenas dimensões é totalmente instável (Fig.4.39). Para amostras de dimensão 10, ultrapassa os 45° nove vezes, atingindo um máximo de 80° : Estabiliza para dimensões maiores, embora para $n = 200$ o ângulo máximo atingido ainda seja $9,32^\circ$.

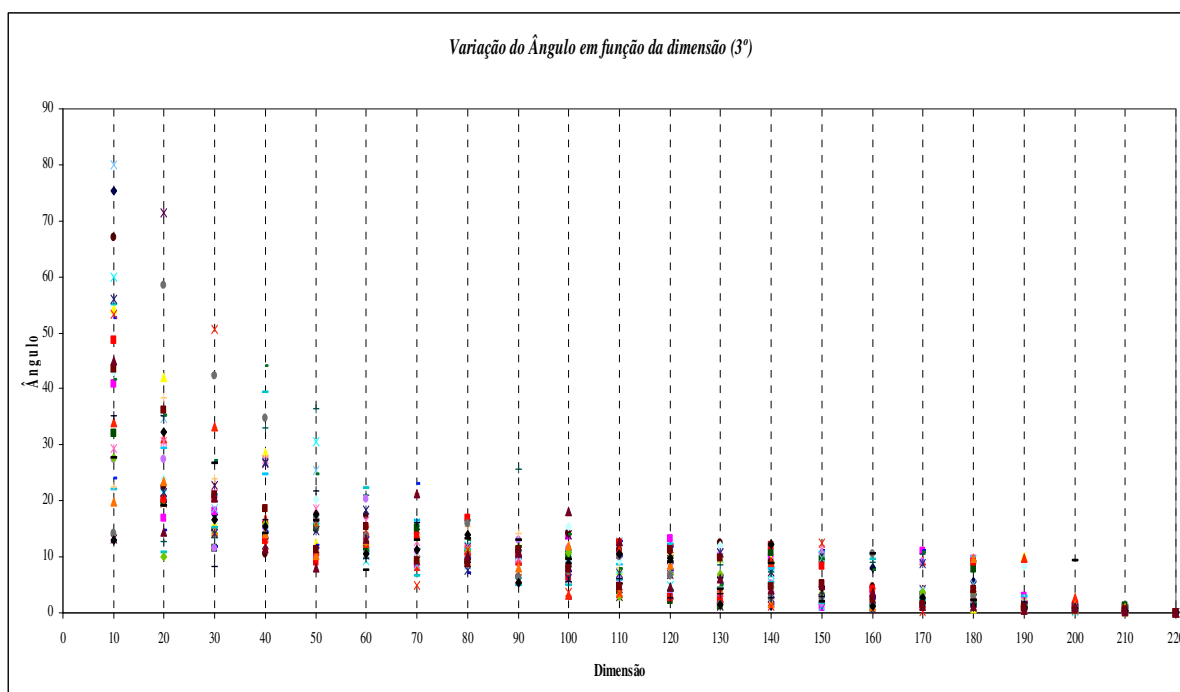


Fig.4.39 – Variação do 3º ângulo em função da dimensão da amostra

O 4º eixo é totalmente instável para amostras pequenas. Note-se que para dimensão 10, o ângulo entre o 4º eixo real e o 4º eixo amostral atinge um máximo de 78°. Aliás o limiar de estabilidade (45°) é ultrapassado 11 vezes. Para dimensão 20 é ultrapassado 4 vezes e para dimensão 30 uma vez.

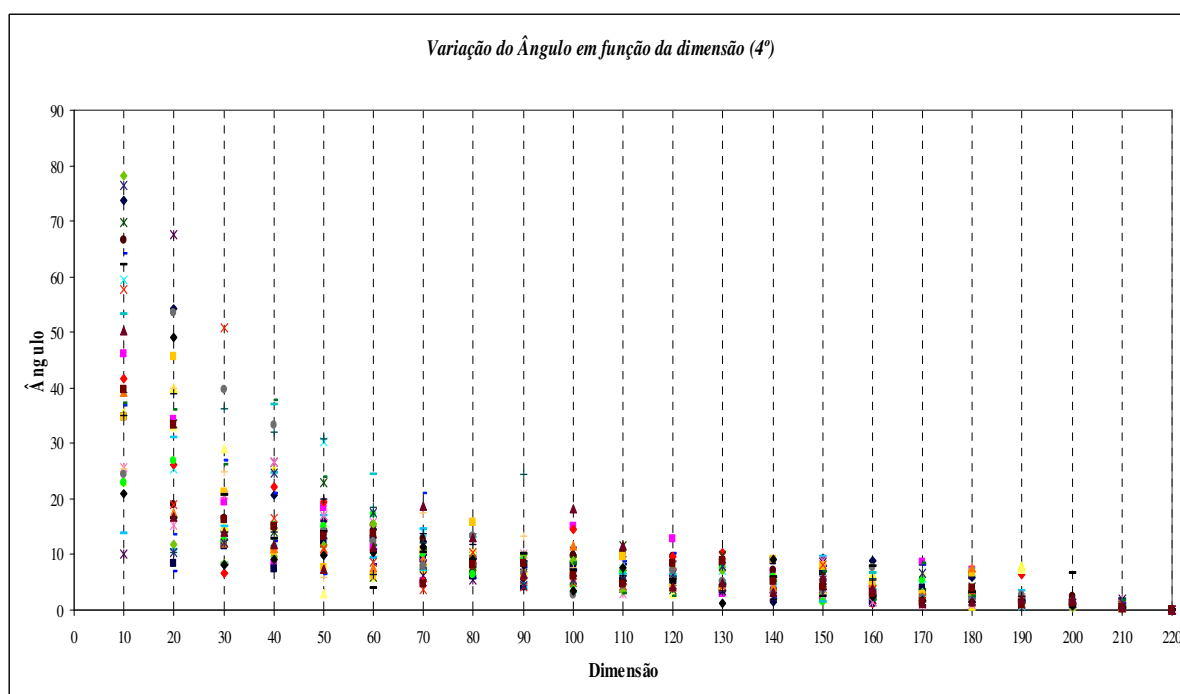


Fig.4.40 – Variação do 4º ângulo em função da dimensão da amostra

Para dimensões maiores o sistema estabiliza: a partir de $n = 130$ o ângulo máximo atingido nunca excede os 10° .

Na figura 4.41 confronta-se a estabilidade dos ângulos (médios) obtidos em relação aos diferentes eixos no caso contínuo (a escuro) e para o caso discreto.

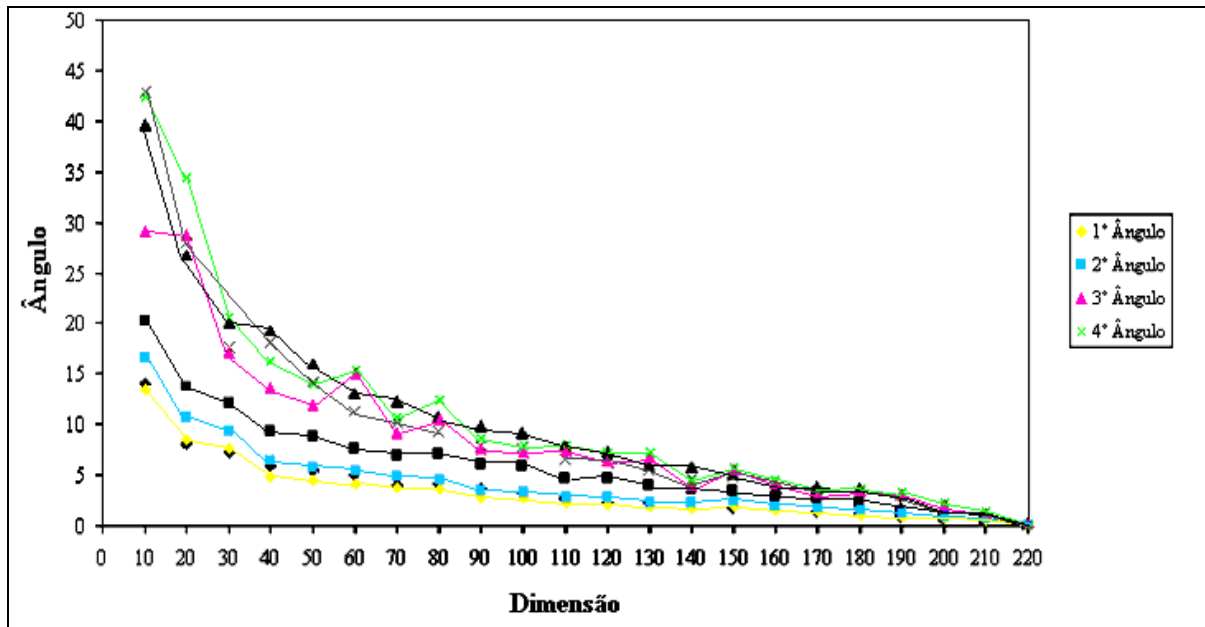


Fig.4.41 – Estabilidade Comparada

Parece, pese embora a similitude da situação discreta e contínua, que para o 2º e 3º ângulos a variação do ângulo é maior no caso contínuo.

4.4 Classificações de Alunos

Consideram-se as classificações em nove disciplinas de cinquenta alunos do 7º ano da Escola Secundária Abade de Baçal de Bragança (Anexo VII). Estes dados referem-se ao 3º período do ano lectivo 1997/98 e as classificações assumem valores entre 1 e 5, inclusive.

Como variáveis tomaram-se apenas as respectivas disciplinas, que se indicam na tabela abaixo:

X_1	→	<i>Português</i>
X_2	→	<i>Educação Física</i>
X_3	→	<i>Inglês</i>
X_4	→	<i>Matemática</i>
X_5	→	<i>Ciências da Natureza</i>
X_6	→	<i>Geografia</i>
X_7	→	<i>História</i>
X_8	→	<i>Francês</i>
X_9	→	<i>Educação Visual</i>

Tabela 4.13

A análise da Tabela 4.14 revela que, em média, as melhores classificações são obtidas para as disciplinas de Educação Visual, X_9 , Educação Física, X_2 , e Geografia, X_6 . Note-se que para Educação Física e Educação Visual a classificação mínima obtida é positiva.

As piores classificações correspondem a cadeiras nucleares: Português, X_1 , Matemática, X_4 , e Ciências da Natureza, X_5 .

	X1	X2	X3	X4	X5	X6	X7	X8	X9
Média	3,24	4,1	3,58	3,28	3,44	4,08	3,58	3,8	4,2
Mediana	3	4	3	3	3	4	3	4	4
Moda	4	4	3	3	4	5	3	3	5
Variância	0,9616	0,5	1,0241	0,9812	1,2310	0,8914	0,7792	0,6939	0,5714
Desvio-padrão *	0,9806	0,7071	1,012	0,9906	1,1095	0,9442	0,8827	0,8329	0,7559
Assimetria	0,1614	-0,1399	0,0782	0,1806	0,0622	-0,3064	0,2042	0,1712	-0,3436
Curtose **	1,9471	2,0396	1,8852	1,9799	1,6876	1,5284	2,2116	1,8858	1,8469
Mínimo	2	3	2	2	2	2	2	2	3
Máximo	5	5	5	5	5	5	5	5	5
Coeficiente de Variação	0,3027	0,1725	0,2827	0,3020	0,3225	0,2314	0,2467	0,2192	0,1799

* - Desvio Padrão Corrigido

** - Definição usada: $k = \frac{E(x-\mu)^4}{\sigma^4}$

Tabela 4.14

As medidas de dispersão (Desvio Padrão, Variância e Coeficiente de Variação), bem como a análise do diagrama caixa de bigodes (Fig.4.42) mostram que existe, como seria de esperar pela dimensão da escala classificatória, pouca dispersão em todas as variáveis.

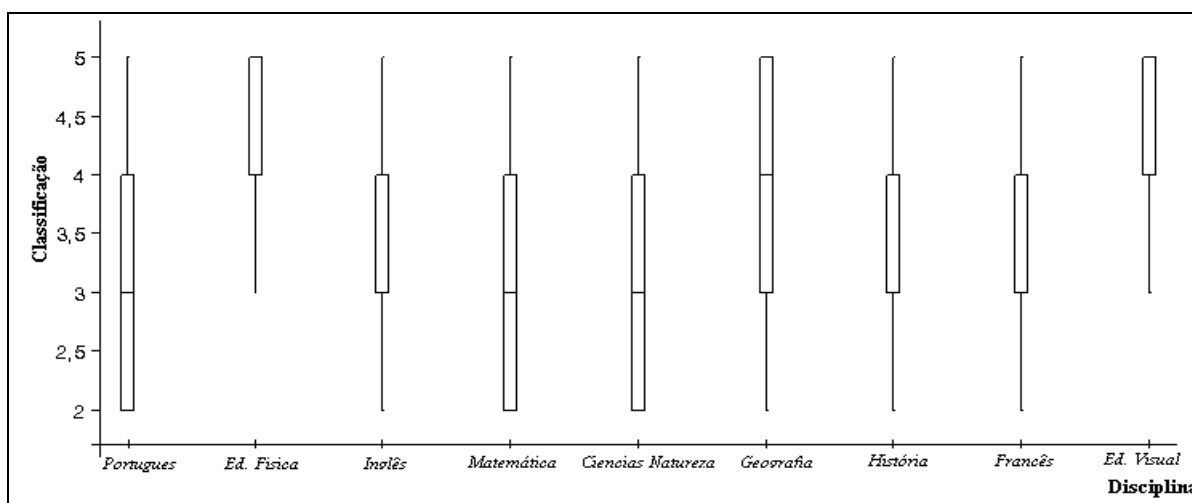


Fig. 4.42 - Diagrama Caixa de Bigodes

Observe-se como para Português, Matemática e Ciências da Natureza 50% das classificações tomam valor entre 2 e 3.

Para as variáveis consideradas obteve-se a seguinte matriz de correlações:

	X1	X2	X3	X4	X5	X6	X7	X8	X9
X1	1	0,49	0,72	0,77	0,80	0,77	0,80	0,81	0,62
X2	0,49	1	0,60	0,43	0,49	0,48	0,49	0,49	0,42
X3	0,72	0,60	1	0,75	0,82	0,63	0,69	0,77	0,49
X4	0,77	0,43	0,75	1	0,81	0,70	0,74	0,79	0,55
X5	0,80	0,49	0,82	0,81	1	0,80	0,84	0,80	0,53
X6	0,77	0,48	0,63	0,70	0,80	1	0,75	0,72	0,52
X7	0,80	0,49	0,69	0,74	0,84	0,75	1	0,77	0,59
X8	0,81	0,49	0,77	0,79	0,80	0,72	0,77	1	0,65
X9	0,62	0,42	0,49	0,55	0,53	0,52	0,59	0,65	1

Tabela 4.15

Organizando a matriz por blocos (de forma que as correlações existentes em cada bloco sejam todas superiores a 50%) verifica-se que se correlacionam fortemente todas as disciplinas com exceção de Educação Física, X_2 , e Educação Visual, X_9 (Tabela 4.16):

	X8	X1	X5	X4	X3	X7	X6	X9	X2
X8	1	0,81	0,8	0,79	0,77	0,77	0,72	0,65	0,49
X1	0,81	1	0,8	0,77	0,72	0,8	0,77	0,62	0,49
X5	0,8	0,8	1	0,81	0,82	0,84	0,8	0,53	0,49
X4	0,79	0,77	0,81	1	0,75	0,74	0,7	0,55	0,43
X3	0,77	0,72	0,82	0,75	1	0,69	0,63	0,49	0,6
X7	0,77	0,8	0,84	0,74	0,69	1	0,75	0,59	0,49
X6	0,72	0,77	0,8	0,7	0,63	0,75	1	0,52	0,48
X9	0,65	0,62	0,53	0,55	0,49	0,59	0,52	1	0,42
X2	0,49	0,49	0,49	0,43	0,6	0,49	0,48	0,42	1

Tabela 4.16

Como resultado, no bloco maior quase todas as correlações são superiores a 70% com exceção do Inglês, X_3 , que apresenta as menores correlações com Geografia (63%), X_6 , e História (69%), X_7 .

No bloco antidiagonal as correlações são quase todas inferiores a 60%, excluindo a correlação de Educação Visual, X_9 , com Francês, X_8 , e Português, X_1 .

Na tabela 4.17 apresentam-se os valores próprios da matriz de correlações:

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
6,39	70,98%	70,98
0,70	7,78%	78,76
0,61	6,73%	85,49
0,39	4,36%	89,85
0,23	2,61%	92,46
0,21	2,35%	94,81
0,21	2,29%	97,10
0,17	1,88%	98,98
0,09	1,02%	100,00%

Tabela 4.17

Apenas a primeira componente explica 71% da inércia total. A escolha do número de componentes a reter poderá ser de duas, já que explicam cerca de 79 % da informação.

Segue-se a representação dos alunos no primeiro plano principal (Fig. 4.43), onde cada aluno é representado pela respectiva classificação em cada uma das disciplinas consideradas:

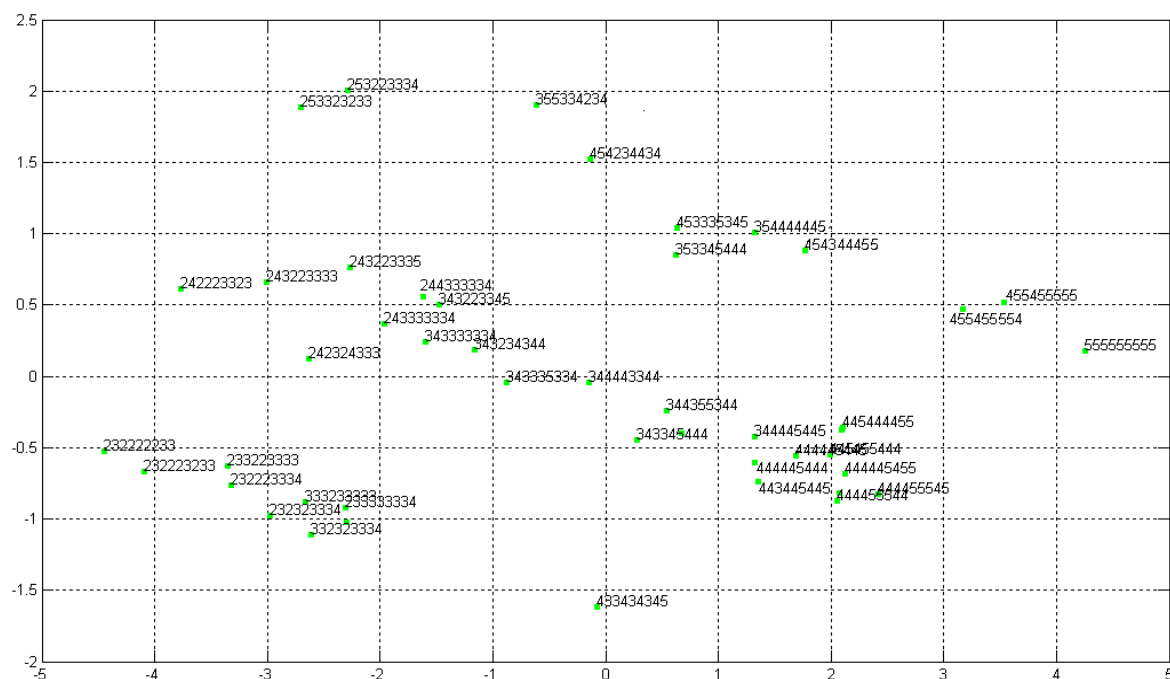


Fig. 4.43 - Indivíduos representados no plano principal (1,2)

O primeiro eixo parece colocar em oposição alunos com boas classificações (4 e 5) em todas as disciplinas e alunos que têm classificações menos boas (2 e 3). No entanto, os alunos com classificação menos boas, têm como denominador comum o facto de todos terem nota positiva a Educação Física e Educação Visual:

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Português	Edu. Física	Inglês	Matemática	C. da Natureza	Geografia	História	Francês	Edu. Visual
5	5	5	5	5	5	5	5	5
4	5	5	4	5	5	5	5	5
4	5	5	4	5	5	5	5	4

vs

2	3	2	2	2	2	2	3	3
2	3	2	2	2	3	2	3	3
2	4	2	2	2	3	3	2	3

Tabela 4.18

Aliás, observando o círculo de correlações (Fig. 4.44) observa-se que as variáveis que mais contribuem para a formação da primeira componente são $X_3, X_4, X_5, X_6, X_7, X_8$, em particular X_5 (Ciências da Natureza). X_2 (Edu. Física) e X_9 (Edu. Visual) apresentam o menor contributo.

	Y1	Y2
X1	0,9056	-0,1053
X2	0,6246	0,7650
X3	0,8598	0,1560
X4	0,8741	-0,1796
X5	0,9231	-0,1175
X6	0,8496	-0,1130
X7	0,8907	-0,1062
X8	0,9057	-0,0914
X9	0,6981	0,0318

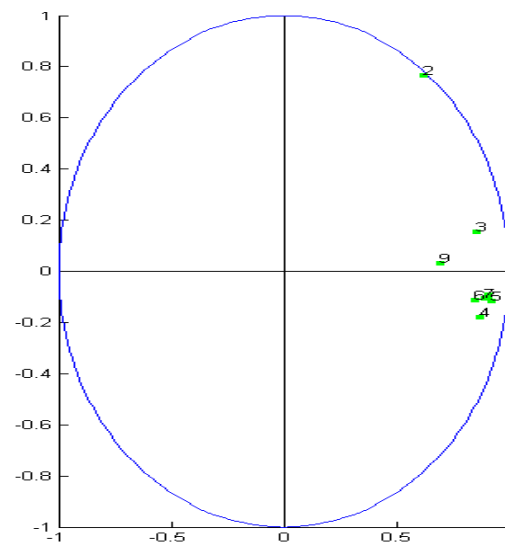


Fig.4.44 – Círculo de Correlações (12)

Portanto, para a formação da primeira componente as variáveis que mais contribuem são as pertencentes ao bloco de maior dimensão na matriz de correlações.

O 2º eixo coloca em oposição os alunos com as seguintes classificações:

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Português	Edu. Física	Inglês	Matemática	C. da Natureza	Geografia	História	Francês	Edu. Visual
2	5	3	2	2	3	3	3	4
2	5	3	3	2	3	2	3	3
3	5	5	3	3	4	2	3	4

vs

4	3	3	4	3	4	3	4	5
---	---	---	---	---	---	---	---	---

Tabela 4.19

O círculo de correlações (Fig. 4.44) mostra que a variável com maior preponderância é X_2 (Edu. Física). De facto, atentando nas oposições verifica-se que se opõem alunos com classificação 5 nesta disciplina a alunos com classificação 3.

Existem, neste eixo, outras oposições, talvez menos evidentes.

Opõem-se alunos com más classificações a Matemática, Português e Ciências da Natureza (grupo de alunos com 5 a Edu. Física) a alunos com boas classificações nestas disciplinas.

Embora sem interesse em termos de inércia explicada, analisa-se a interpretação da terceira componente:

O 3º eixo (Fig. 4.45) coloca em oposição os alunos com as seguintes classificações:

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Português	Edu. Física	Inglês	Matemática	C. da Natureza	Geografia	História	Francês	Edu. Visual
3	4	3	2	2	3	3	4	5
2	4	3	2	2	3	3	3	5

vs

4	4	5	5	4	5	4	5	3
---	---	---	---	---	---	---	---	---

Tabela 4.20

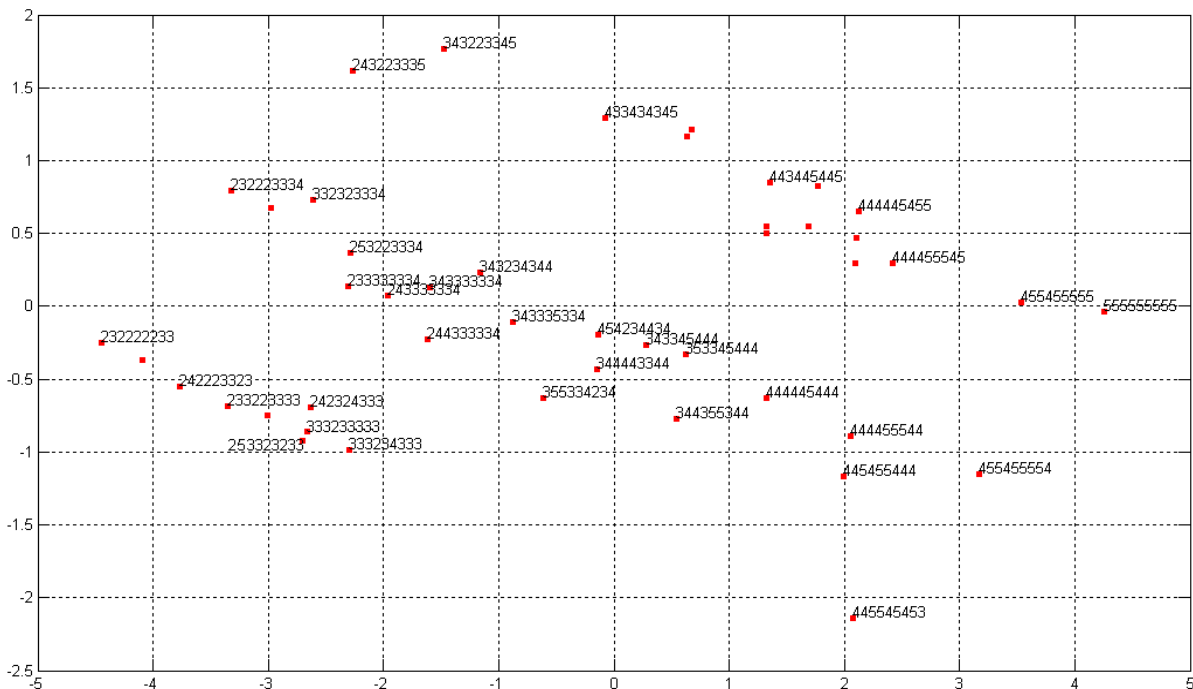


Fig. 4.45 - Indivíduos representados no plano principal (1,3)

Opõem-se alunos com classificação baixa a Português, Inglês, Matemática, Ciências da Natureza, Geografia e História e com classificação alta a Educação Visual a alunos com boa classificação nas primeiras disciplinas e classificação baixa a Educação Visual.

O círculo de correlações (Fig. 4.46) apresenta que a variável com maior contributo é a variável X_9 , Educação Visual.

	Y3
X1	0,0410
X2	-0,0357
X3	-0,2362
X4	-0,0953
X5	-0,2021
X6	-0,0862
X7	-0,0145
X8	0,0640
X9	0,6965

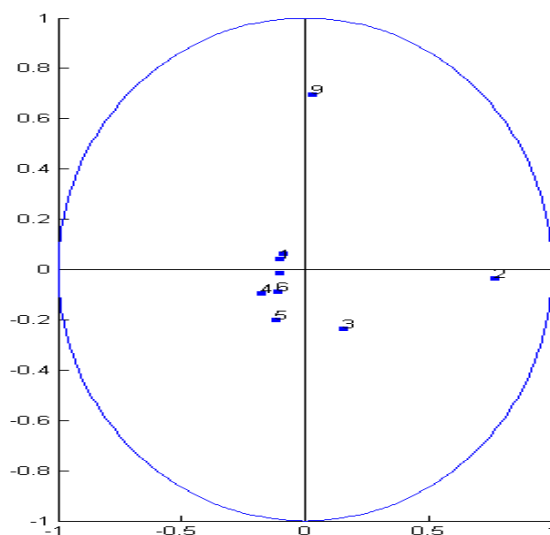


Fig. 4.46 - Círculo de Correlações (23)

Assim, o bloco de variáveis de correlações mais fortes está associado ao primeiro eixo, e cada uma das restantes variáveis isoladas, Educação Física e Educação Visual, associam-se ao 2º e 3º eixo, respectivamente.

Passam agora a considerar-se os valores das notas ordenados por ordem crescente, para cada aluno. Ou seja, as variáveis deixam de corresponder às disciplinas e passamos a ter:

X_1	→	<i>Classificação Mínima</i>
$X_2 \dots X_8$	→	<i>Classificações Intermédias</i>
X_9	→	<i>Classificação Máxima</i>

Tabela 4.21

Obviamente, que para as variáveis representativas das menores classificações há uma redução do valor médio e um aumento para as variáveis representativas das melhores classificações (Tabela 4.22).

Naturalmente que a moda para a variável X_1 é 2 e 5 para X_9 .

O facto de se ordenarem as notas não significa que para a variável X_1 apenas existam valores baixos e altos para X_9 . De facto, o valor máximo de X_1 é 5 e 3 é o mínimo para X_9 . Isto, porque há alunos com boas classificações a todas as disciplinas e outros com classificações baixas.

	X1	X2	X3	X4	X5	X6	X7	X8	X9
Média	2,96	3,24	3,3	3,58	3,64	3,78	3,98	4,24	4,58
Mediana	3	3	3	4	4	4	4	4	5
Moda	2	3	3	3	4	4	4	5	5
Variância	1,0188	0,9208	0,9082	0,6976	0,6433	0,6241	0,6731	0,6759	0,4118
Desvio-padrão	1,0093	0,9596	0,9529	0,8352	0,8020	0,7899	0,8204	0,8221	0,6418
Assimetria	0,6815	0,2058	0,2287	0,0627	0,0149	0,1554	0,0366	-0,4651	-1,2378
Curtose	2,3066	2,0574	2,1473	2,4071	2,4755	2,0998	1,517	1,6603	3,3674
Mínimo	2	2	2	2	2	2	3	3	3
Máximo	5	5	5	5	5	5	5	5	5
Coefficiente de Variação	0,3409	0,2962	0,2888	0,2333	0,2203	0,2089	0,2061	0,1939	0,14012

Tabela 4.22

Continuam a verificar-se baixos valores de dispersão.

Segue-se a matriz de correlações para as variáveis consideradas:

	X1	X2	X3	X4	X5	X6	X7	X8	X9
X1	1	0,90	0,88	0,85	0,86	0,81	0,81	0,73	0,57
X2	0,90	1	0,97	0,87	0,86	0,85	0,84	0,78	0,56
X3	0,88	0,97	1	0,88	0,87	0,85	0,84	0,77	0,58
X4	0,85	0,87	0,88	1	0,96	0,88	0,79	0,77	0,65
X5	0,86	0,86	0,87	0,96	1	0,90	0,83	0,78	0,69
X6	0,81	0,85	0,85	0,88	0,90	1	0,87	0,81	0,66
X7	0,81	0,84	0,84	0,79	0,83	0,87	1	0,85	0,68
X8	0,73	0,78	0,77	0,77	0,78	0,81	0,85	1	0,78
X9	0,57	0,56	0,58	0,65	0,69	0,66	0,68	0,78	1

Tabela 4.23

Em relação à matriz anterior verifica-se que as correlações aumentam. Diminui apenas a correlação entre X_1 (mínimo) e X_9 (máximo). Pode observar-se que todas as correlações são superiores a 50% e que se organizam por ordem decrescente.

Na tabela seguinte estão apresentados os valores próprios obtidos para a matriz de correlações:

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
7,44	82,62%	82,62%
0,65	7,17%	89,79%
0,29	3,19%	92,99%
0,21	2,36%	95,35%
0,15	1,67%	97,02%
0,13	1,41%	98,42%
0,08	0,90%	99,32%
0,03	0,39%	99,71%
0,03	0,29%	100,00%

Tabela 4.24

Neste caso, apenas duas componentes explicam 89,79% de inércia, ou seja, mais inércia que a anteriormente explicada por três componentes.

Após a ordenação das notas que tipo de interpretação passam a ter as componentes principais?

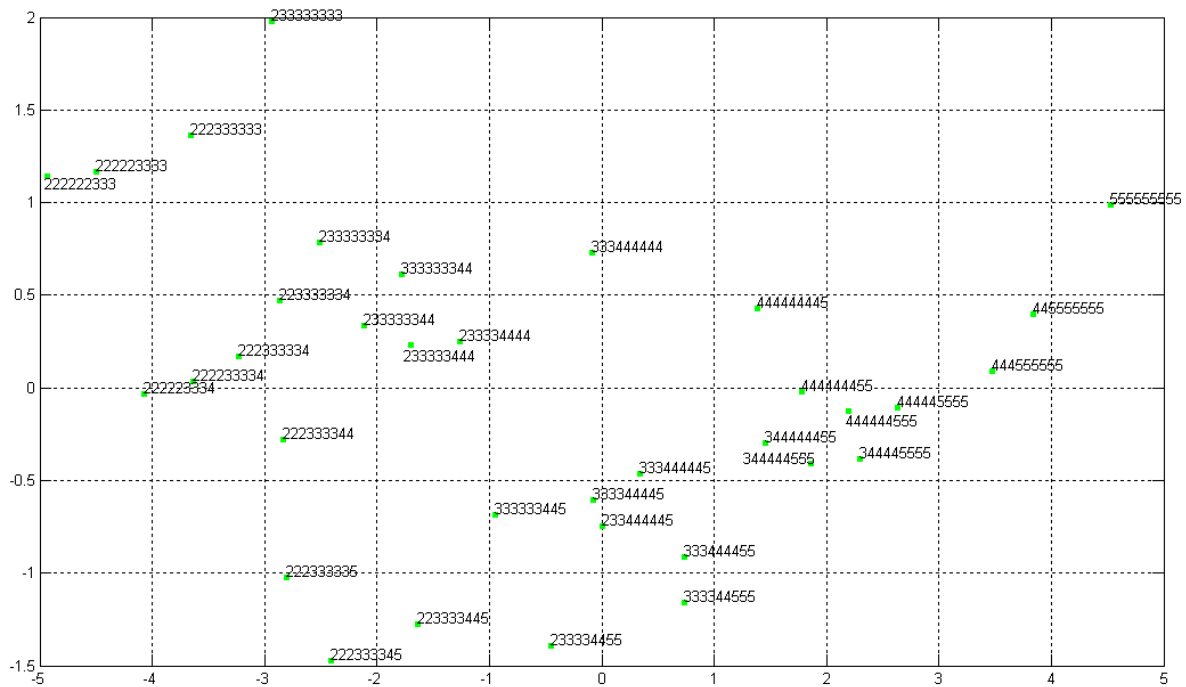


Fig. 4.47- Indivíduos representados no plano principal (1,2)

A representação dos indivíduos (Fig. 4.47) no plano principal (1,2) mostra que se opõem os seguintes indivíduos:

5	5	5	5	5	5	5	5	5
4	4	5	5	5	5	5	5	5
4	4	4	5	5	5	5	5	5

vs

2	2	2	2	2	2	3	3	3
2	2	2	2	2	3	3	3	3
2	2	2	2	2	3	3	3	4

Tabela 4.25

Ou seja, encontram-se em oposição alunos com todas as classificações elevadas a alunos com todas as classificações baixas.

O círculo de correlações (Fig. 4.48) mostra que todas as variáveis se correlacionam fortemente com a primeira componente. A variável X_9 (máximo), embora muito correlacionada, é aquela que menos contribui.

	Y1	Y2
X1	0,9105	0,2255
X2	0,9370	0,2413
X3	0,9382	0,2320
X4	0,9394	0,0918
X5	0,9504	0,0409
X6	0,9358	0,0129
X7	0,9208	-0,0701
X8	0,8855	-0,2966
X9	0,7450	-0,6161

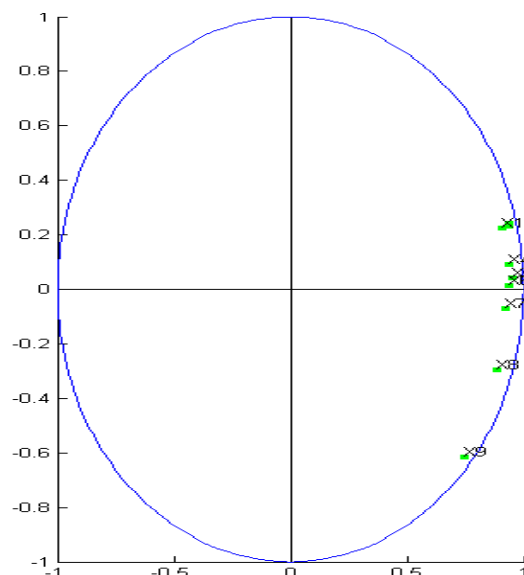


Fig. 4.48- Círculo de Correlações (12)

O segundo eixo opõe:

2	3	3	3	3	3	3	3	3
---	---	---	---	---	---	---	---	---

vs

2	2	2	3	3	3	3	4	5
2	3	3	3	3	4	4	5	5
2	2	3	3	3	3	4	4	5

Tabela 4.26

Parece opor indivíduos que tem todas as classificações médias/baixas a indivíduos com pelo menos duas classificações elevadas e as restantes médias baixas.

Daí que as variáveis que mais contribuem para a formação da segunda componente sejam as variáveis X_8 e X_9 , uma vez que o que distingue estes indivíduos são, precisamente, as duas notas mais elevada

Quer no caso ordenado, quer no caso desordenado é possível fazer uma distinção semelhante dos indivíduos. No entanto, a partir do momento em que se ordenam as classificações perde-se a informação relativa às disciplinas, o que explica o aumento da informação explicada apenas por duas componentes.

Consideram-se agora apenas três disciplinas básicas: Português, Matemática e Inglês e aplicam-se as variáveis usadas nos paralelepípedos (Tabela 4.27).

Numa primeira abordagem não se ordenam as notas, ou seja, tem-se

X_1	→	<i>Matemática</i>
X_2	→	<i>Português</i>
X_3	→	<i>Inglês</i>
X_4	→	<i>Produto das Notas</i>
X_5	→	<i>Soma dos produtos das notas duas a duas</i>
X_6	→	<i>Nota Máxima</i>
X_7	→	<i>Nota Mínima</i>
X_8	→	<i>Nota Média</i>
X_9	→	<i>Média das Produtos das Notas duas a Duas</i>
X_{10}	→	$\frac{\text{Nota Máximo}}{\text{Nota Mínimo}}$
X_{11}	→	<i>Nota Máxima – Nota Mínima</i>
X_{12}	→	$\frac{\text{Mínimo Pr oduto das notas duas a duas}}{\text{Máximo Pr oduto das notas duas a duas}}$

Tabela 4.27

Neste caso, as variáveis não são geométricas pelo que a interpretação não é tão clara.

Obtém-se a seguinte matriz de correlações:

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1,00	0,77	0,75	0,91	0,92	0,82	0,90	0,92	0,92	-0,41	-0,24	0,39
X2	0,77	1,00	0,72	0,88	0,90	0,79	0,89	0,91	0,90	-0,40	-0,26	0,40
X3	0,75	0,72	1,00	0,86	0,89	0,94	0,80	0,91	0,89	-0,09	0,10	0,10
X4	0,91	0,88	0,86	1,00	0,99	0,87	0,96	0,97	0,99	-0,42	-0,27	0,42
X5	0,92	0,90	0,89	0,99	1,00	0,91	0,96	0,99	1,00	-0,38	-0,21	0,38
X6	0,82	0,79	0,94	0,87	0,91	1,00	0,81	0,93	0,91	-0,03	0,18	0,02
X7	0,90	0,89	0,80	0,96	0,96	0,81	1,00	0,94	0,96	-0,59	-0,43	0,59
X8	0,92	0,91	0,91	0,97	0,99	0,93	0,94	1,00	0,99	-0,32	-0,14	0,32
X9	0,92	0,90	0,89	0,99	1,00	0,91	0,96	0,99	1,00	-0,38	-0,21	0,38
X10	-0,41	-0,40	-0,09	-0,42	-0,38	-0,03	-0,59	-0,32	-0,38	1,00	0,94	-0,98
X11	-0,24	-0,26	0,10	-0,27	-0,21	0,18	-0,43	-0,14	-0,21	0,94	1,00	-0,95
X12	0,39	0,40	0,10	0,42	0,38	0,02	0,59	0,32	0,38	-0,98	-0,95	1,00

Tabela 4.28

Organizando a matriz de correlações em blocos de fortes correlações (Tabela 4.29) obtêm-se apenas dois blocos: um constituído pelas variáveis função dos extremos e o de maior dimensão composto pelas restantes variáveis (analíticas, cada uma das classificações, máximo e mínimo):

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1,00	0,77	0,75	0,91	0,92	0,82	0,90	0,92	0,92	-0,41	-0,24	0,39
X2	0,77	1,00	0,72	0,88	0,90	0,79	0,89	0,91	0,90	-0,40	-0,26	0,40
X3	0,75	0,72	1,00	0,86	0,89	0,94	0,80	0,91	0,89	-0,09	0,10	0,10
X4	0,91	0,88	0,86	1,00	0,99	0,87	0,96	0,97	0,99	-0,42	-0,27	0,42
X5	0,92	0,90	0,89	0,99	1,00	0,91	0,96	0,99	1,00	-0,38	-0,21	0,38
X6	0,82	0,79	0,94	0,87	0,91	1,00	0,81	0,93	0,91	-0,03	0,18	0,02
X7	0,90	0,89	0,80	0,96	0,96	0,81	1,00	0,94	0,96	-0,59	-0,43	0,59
X8	0,92	0,91	0,91	0,97	0,99	0,93	0,94	1,00	0,99	-0,32	-0,14	0,32
X9	0,92	0,90	0,89	0,99	1,00	0,91	0,96	0,99	1,00	-0,38	-0,21	0,38
X10	-0,41	-0,40	-0,09	-0,42	-0,38	-0,03	-0,59	-0,32	-0,38	1,00	0,94	-0,98
X11	-0,24	-0,26	0,10	-0,27	-0,21	0,18	-0,43	-0,14	-0,21	0,94	1,00	-0,95
X12	0,39	0,40	0,10	0,42	0,38	0,02	0,59	0,32	0,38	-0,98	-0,95	1,00

Tabela 4.29

Neste caso, duas componentes explicam 98% da inércia total, sendo que apenas a primeira componente contribui com 77,1 % da informação (Tabela 4.30).

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
8,6035	71,70	71,70
2,8226	23,52	95,22
0,23805	1,98	97,20
0,1937	1,61	98,82
0,075702	0,63	99,45
0,028172	0,23	99,68
2,21E-02	0,18	99,87
1,61E-02	0,13	100,00
6,35E-05	0,00	100,00
3,22E-16	0,00	100,00
7,97E-17	0,00	100,00
-9,26E-16	0,00	100,00

Tabela 4.30

Na figura 4.49 representam-se os indivíduos no primeiro plano principal:

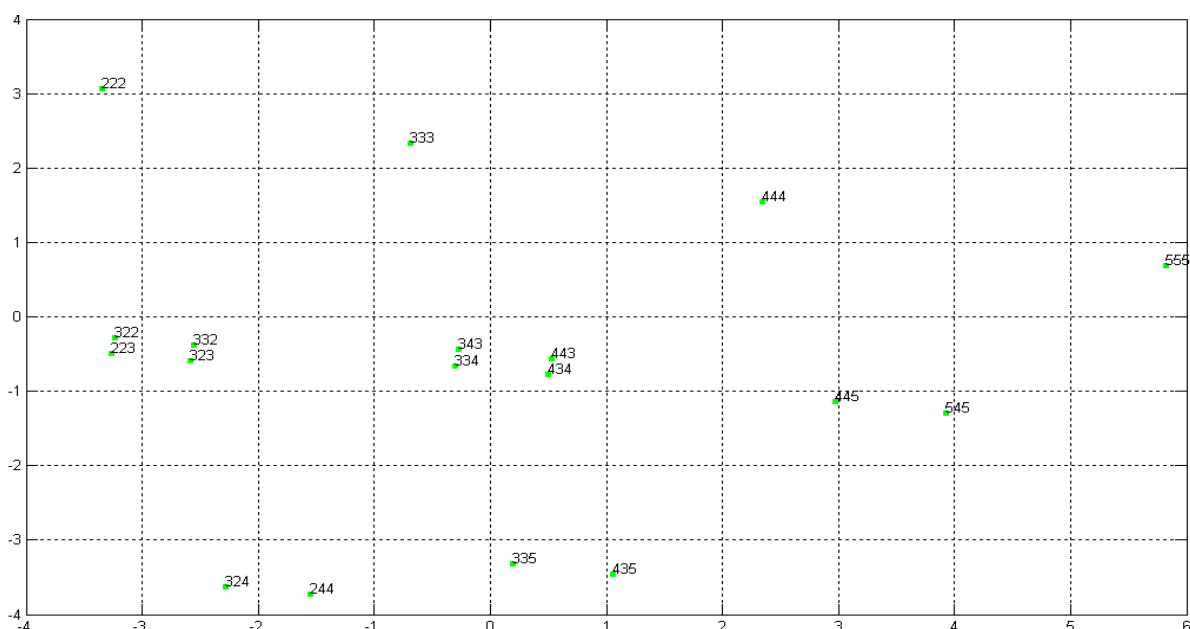


Fig.4.49 – Indivíduos representados no plano principal (1,2)

O primeiro eixo opõe alunos com notas altas em todas as disciplinas a alunos com notas baixas em pelo menos duas dessas disciplinas. O equivalente, em termos geométricos, à oposição entre grandes e pequenos volumes/áreas.

O segundo eixo opõe alunos com todas as notas iguais (independentemente do valor) a alunos com duas classificações muito diferentes. Ou seja, opõe-se uma diferença pequena (neste caso, nula) entre Máximo e Mínimo em oposição a uma grande diferença entre estes valores.

Olhando ao círculo de correlações (Fig. 4.50), tem-se que as variáveis que mais contribuem para a formação da primeira componente são as variáveis pertencentes ao primeiro bloco de correlações. Para a segunda componente são as variáveis função dos extremos, pertencentes ao segundo maior bloco, que mais contribuem para a sua formação.

	Y1	Y2
X1	0,9244	-0,0455
X2	0,9090	-0,0255
X3	0,8548	-0,3977
X4	0,9839	-0,0519
X5	0,9917	-0,1088
X6	0,8696	-0,4728
X7	0,9858	0,1305
X8	0,9824	-0,1737
X9	0,9917	-0,1088
X10	-0,4824	-0,8598
X11	-0,3128	-0,9384
X12	0,4831	0,8630

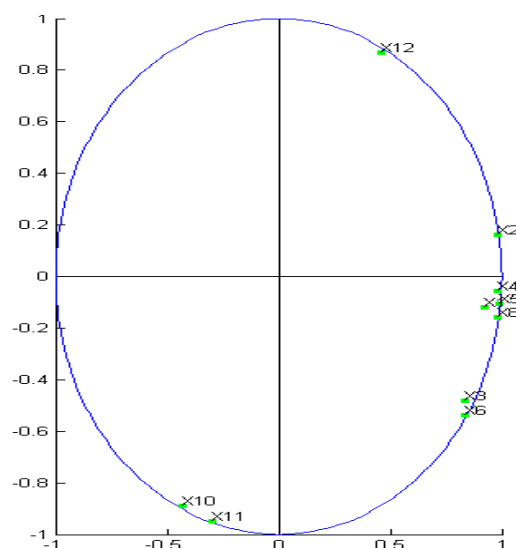


Fig.4.50 – Círculo de Correlações (12)

As disciplinas propriamente ditas não têm intervenção na distinção dos indivíduos. Provavelmente, por todas serem disciplinas nucleares o comportamento dos alunos é semelhante.

Tomando as classificações ordenadas, redefinem-se as variáveis iniciais do seguinte modo:

X_1	→	<i>Nota Mínima</i>
X_2	→	<i>Nota ‘Intermédia’</i>
X_3	→	<i>Nota Máxima</i>
X_4	→	<i>Produto das Notas</i>
X_5	→	<i>Soma dos produtos das notas duas a duas</i>
X_6	→	<i>Nota Máxima</i>
X_7	→	<i>Nota Mínima</i>
X_8	→	<i>Nota Média</i>
X_9	→	<i>Média das Produtos das Notas duas a Duas</i>
X_{10}	→	$\frac{\text{Nota Máximo}}{\text{Nota Mínimo}}$
X_{11}	→	<i>Nota Máxima – Nota Mínima</i>
X_{12}	→	$\frac{\text{Mínimo Pr oduto das notas duas a duas}}{\text{Máximo Pr oduto das notas duas a duas}}$

Tabela 4.31

Para esta situação obtém-se a seguinte matriz de correlações:

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1,00	0,85	0,81	0,96	0,96	0,81	1,00	0,94	0,96	-0,59	-0,43	0,59
X2	0,85	1,00	0,85	0,91	0,94	0,85	0,85	0,95	0,94	-0,27	-0,13	0,28
X3	0,81	0,85	1,00	0,87	0,91	1,00	0,81	0,93	0,91	-0,03	0,18	0,02
X4	0,96	0,91	0,87	1,00	0,99	0,87	0,96	0,97	0,99	-0,42	-0,27	0,42
X5	0,96	0,94	0,91	0,99	1,00	0,91	0,96	0,99	1,00	-0,38	-0,21	0,38
X6	0,81	0,85	1,00	0,87	0,91	1,00	0,81	0,93	0,91	-0,03	0,18	0,02
X7	1,00	0,85	0,81	0,96	0,96	0,81	1,00	0,94	0,96	-0,59	-0,43	0,59
X8	0,94	0,95	0,93	0,97	0,99	0,93	0,94	1,00	0,99	-0,32	-0,14	0,32
X9	0,96	0,94	0,91	0,99	1,00	0,91	0,96	0,99	1,00	-0,38	-0,21	0,38
X10	-0,59	-0,27	-0,03	-0,42	-0,38	-0,03	-0,59	-0,32	-0,38	1,00	0,94	-0,98
X11	-0,43	-0,13	0,18	-0,27	-0,21	0,18	-0,43	-0,14	-0,21	0,94	1,00	-0,95
X12	0,59	0,28	0,02	0,42	0,38	0,02	0,59	0,32	0,38	-0,98	-0,95	1,00

Tabela 4.32

Genericamente, todas as correlações se mantêm e continuam a poder considerar-se apenas dois blocos de fortes correlações.

Verifica-se uma muito ligeira diminuição (Tabela 4.33) da informação explicada pelo primeiro eixo, embora no total as duas primeiras componentes continuem a explicar cerca de 98% da informação total.

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
8,7684	73,07	73,07
2,9391	24,49	97,56
0,17643	1,47	99,03
0,068927	0,57	99,61
0,028004	0,23	99,84
0,019095	0,16	100,00
6,56E-05	0,00	100,00
2,45E-16	0,00	100,00
1,11E-16	0,00	100,00
3,63E-17	0,00	100,00
-5,20E-17	0,00	100,00
-1,83E-15	0,00	100,00

Tabela 4.33

A representação dos indivíduos no primeiro plano principal (Fig.4.51) mostra que se encontram em oposição, tal como no caso não ordenado, notas elevadas (ou seja, grandes volumes/áreas) a notas baixas.

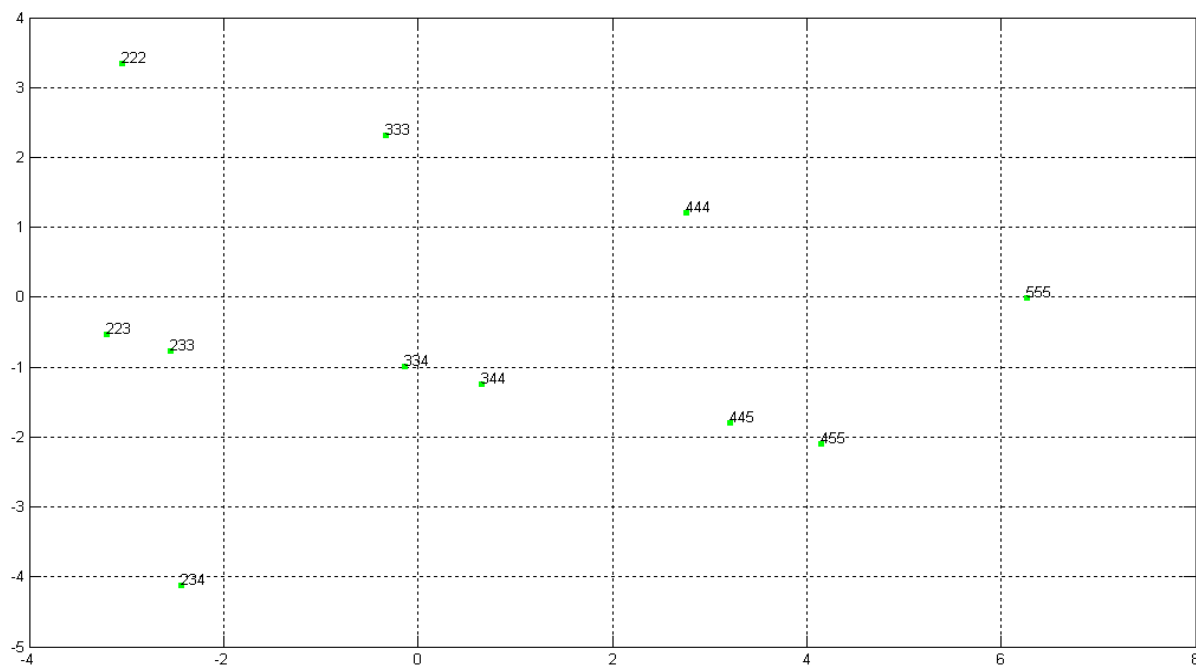


Fig.4.51 – Indivíduos representados no plano principal (1,2)

Contudo verifica-se que a não ordenação fornecia mais informação no sentido que também se podia concluir em relação ao valor em cada disciplina.

Como o objectivo é distinguir os alunos, e o seu comportamento nas disciplinas consideradas é semelhante, o caso desordenado apenas aumenta a inércia sem contribuir para uma distinção mais clara dos alunos.

No círculo de correlações (Fig.4.52), o comportamento é análogo: variáveis do maior bloco com maior contributo para a primeira componente e do menor bloco para a segunda componente:

	Y1	Y2
X1	0,9849	0,1383
X2	0,9222	-0,1750
X3	0,8727	-0,4662
X4	0,9841	-0,0443
X5	0,9924	-0,1022
X6	0,8727	-0,4662
X7	0,9849	0,1383
X8	0,9836	-0,1683
X9	0,9924	-0,1022
X10	-0,4768	-0,8628
X11	-0,3066	-0,9413
X12	0,4776	0,8683

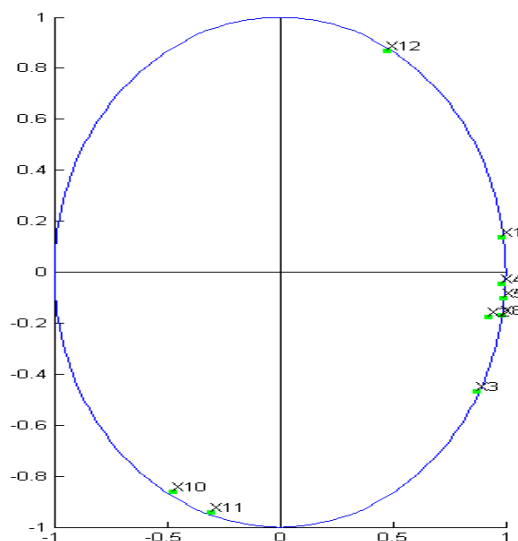


Fig.4.52 – Círculo de Correlações (12)

Para ser mais directamente comparável com o caso inicial seria interessante que as notas variassem de 1 a 10. Tal estudo fica para nova oportunidade.

4.6 Caso Multinormal

Aplicando a matriz de correlações obtida em **3.4.1** (Tabela 3.3) gerou-se uma variável multinormal (220×12) com vector média nulo.

Pode verificar-se que apesar da matriz de correlações ser idêntica a organização dos indivíduos é muito diferente (Fig.4.46 e Fig.4.47), dado que o tipo de valores obtido não traduz a situação geradora da matriz de correlações aplicada.

Obviamente que a representação das variáveis no círculo de correlações é igual:

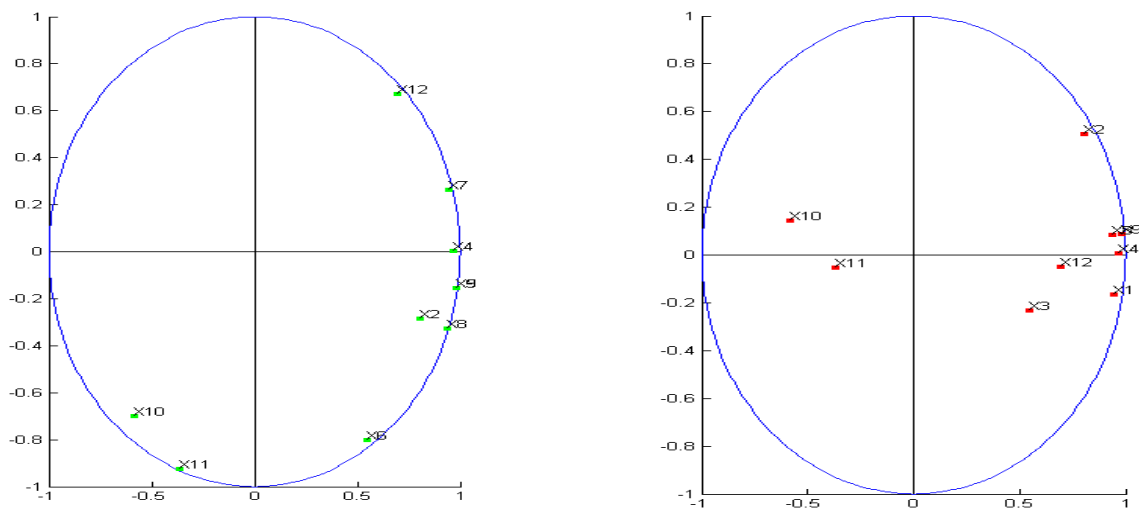


Fig.4.53 – Círculo de Correlações (12) e Círculo de Correlações (13)

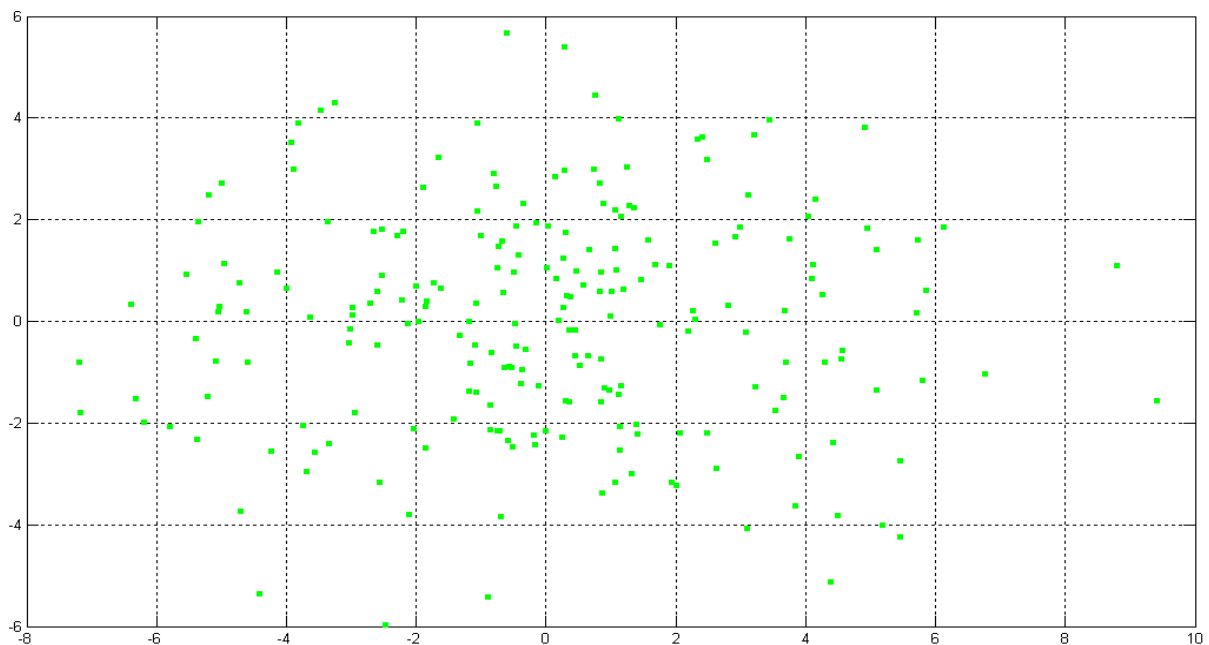


Fig.4.54 – Indivíduos representados no plano principal (1,2)

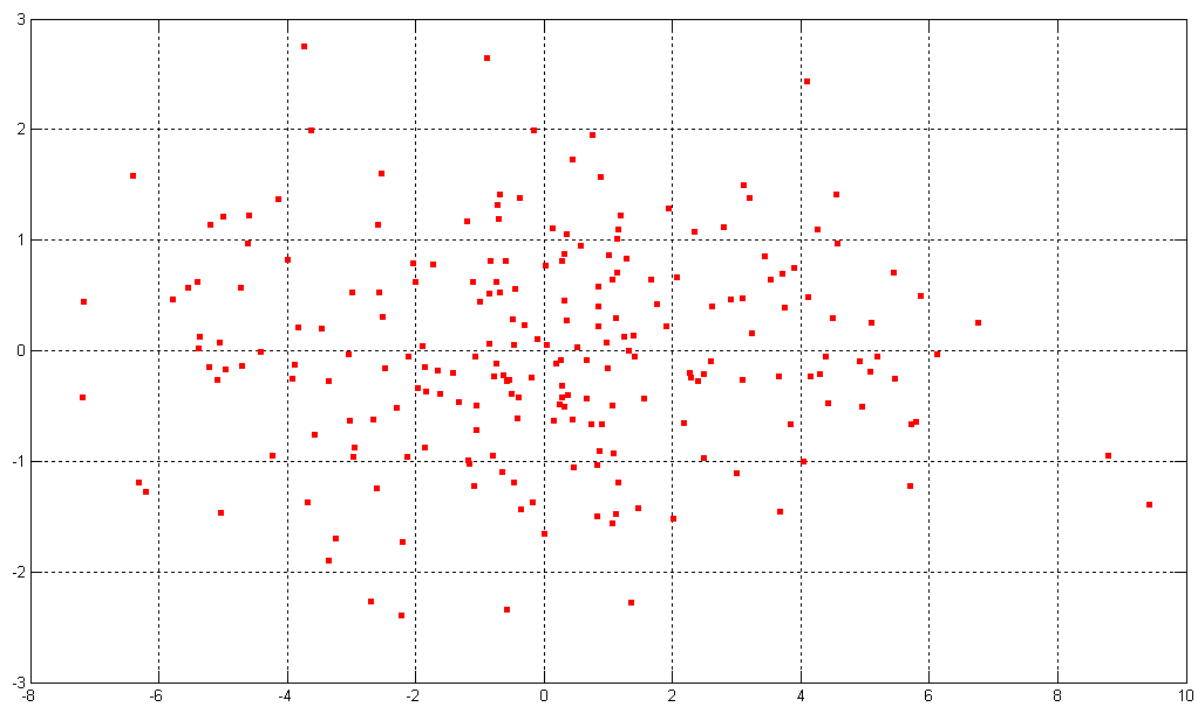


Fig.4.55 – Indivíduos representados no plano principal (1,3)

Como era de prever, não existe qualquer geometria poligonal.

Capítulo V

Variáveis FC-FS

As variáveis consideradas no estudo dos paralelepípedos foram interpretadas na formação do primeiro, segundo e terceiro eixos. Analisaram-se os tipos de variáveis, as correlações e as componentes a que estão ligadas, distinguindo-as.

A pergunta que se agora se põe é se se toma um grande risco de se obter uma interpretação diferente usando outras variáveis.

No sentido de testar a estabilidade do sistema foram introduzidas quatro novas variáveis, consideradas em [1] e não utilizadas no estudo anterior:

X_{13}	<i>Diagonal</i>	→	$\sqrt{X_1^2 + X_2^2 + X_3^2}$
X_{14}	$\frac{\text{Raio da menor esfera circunscrita}}{\text{Raio da maior esfera inscrita}}$	→	$\frac{X_{13}/2}{X_1/2}$
X_{15}	$\frac{\text{Dimensão maior} - \text{Dimensão Intermédia}}{\text{Dimensão Menor}}$	→	$\frac{X_3 - X_2}{X_1}$
X_{16}	$\frac{\text{Área}}{\text{Volume}}$	→	$\frac{X_5}{X_4}$

Tabela 5.1

Note-se que as variáveis introduzidas são claramente mais sofisticadas que as até aqui consideradas, tornando também mais complicada a sua interpretação.

Observando a matriz de correlações (Tabela 5.2) verifica-se que a variável X_{13} se correlaciona fortemente com todas as variáveis iniciais com excepção das variáveis função dos extremos (X_{10} , X_{11} e X_{12}). X_{14} e X_{15} , pelo contrário, apresentam forte correlação com estas variáveis. X_{16} tem correlação significativa com todas as variáveis, excluindo X_{11} .

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16
X1	1	0,58	0,33	0,91	0,86	0,33	1	0,77	0,86	-0,71	-0,58	0,83	0,65	-0,68	-0,67	-0,77
X2	0,58	1	0,58	0,74	0,86	0,58	0,58	0,89	0,86	-0,21	0,00	0,28	0,86	-0,08	-0,02	-0,71
X3	0,33	0,58	1	0,50	0,63	1	0,33	0,77	0,63	0,21	0,58	-0,18	0,88	0,21	0,20	-0,50
X4	0,91	0,74	0,50	1	0,97	0,50	0,91	0,88	0,97	-0,49	-0,35	0,61	0,80	-0,45	-0,44	-0,68
X5	0,86	0,86	0,63	0,97	1	0,63	0,86	0,97	1	-0,42	-0,20	0,52	0,91	-0,36	-0,33	-0,76
X6	0,33	0,58	1	0,50	0,63	1	0,33	0,77	0,63	0,21	0,58	-0,18	0,88	0,21	0,20	-0,50
X7	1	0,58	0,33	0,91	0,86	0,33	1	0,77	0,86	-0,71	-0,58	0,83	0,65	-0,68	-0,67	-0,77
X8	0,77	0,89	0,77	0,88	0,97	0,77	0,77	1	0,97	-0,29	0,00	0,37	0,98	-0,22	-0,19	-0,81
X9	0,86	0,86	0,63	0,97	1	0,63	0,86	0,97	1	-0,42	-0,20	0,52	0,91	-0,36	-0,33	-0,76
X10	-0,71	-0,21	0,21	-0,49	-0,42	0,21	-0,71	-0,29	-0,42	1	0,79	-0,82	-0,10	0,98	0,96	0,61
X11	-0,58	0,00	0,58	-0,35	-0,20	0,58	-0,58	0,00	-0,20	0,79	1	-0,87	0,20	0,77	0,76	0,23
X12	0,83	0,28	-0,18	0,61	0,52	-0,18	0,83	0,37	0,52	-0,82	-0,87	1	0,20	-0,80	-0,79	-0,53
X13	0,65	0,86	0,88	0,80	0,91	0,88	0,65	0,98	0,91	-0,10	0,20	0,20	1	-0,05	-0,02	-0,72
X14	-0,68	-0,08	0,21	-0,45	-0,36	0,21	-0,68	-0,22	-0,36	0,98	0,77	-0,80	-0,05	1	0,99	0,55
X15	-0,67	-0,02	0,20	-0,44	-0,33	0,20	-0,67	-0,19	-0,33	0,96	0,76	-0,79	-0,02	0,99	1	0,51
X16	-0,77	-0,71	-0,50	-0,68	-0,76	-0,50	-0,77	-0,81	-0,76	0,61	0,23	-0,53	-0,72	0,55	0,51	1

Tabela 5.2

Organiza-se a matriz de correlações em blocos de fortes correlações:

	X1	X7	X8	X9	X5	X2	X4	X13	X16	X3	X6	X10	X11	X12	X14	X15
X1	1	1	0,77	0,86	0,86	0,58	0,91	0,65	-0,77	0,33	0,33	-0,71	-0,58	0,83	-0,68	-0,67
X7	1	1	0,77	0,86	0,86	0,58	0,91	0,65	-0,77	0,33	0,33	-0,71	-0,58	0,83	-0,68	-0,67
X8	0,77	0,77	1	0,97	0,97	0,89	0,88	0,98	-0,81	0,77	0,77	-0,29	0	0,37	-0,22	-0,19
X9	0,86	0,86	0,97	1	1	0,86	0,97	0,91	-0,76	0,63	0,63	-0,42	-0,2	0,52	-0,36	-0,33
X5	0,86	0,86	0,97	1	1	0,86	0,97	0,91	-0,76	0,63	0,63	-0,42	-0,2	0,52	-0,36	-0,33
X2	0,58	0,58	0,89	0,86	0,86	1	0,74	0,86	-0,71	0,58	0,58	-0,21	0	0,28	-0,08	-0,02
X4	0,91	0,91	0,88	0,97	0,97	0,74	1	0,8	-0,68	0,5	0,5	-0,49	-0,35	0,61	-0,45	-0,44
X13	0,65	0,65	0,98	0,91	0,91	0,86	0,8	1	-0,72	0,88	0,88	-0,1	0,2	0,2	-0,05	-0,02
X16	-0,77	-0,77	-0,81	-0,76	-0,76	-0,71	-0,68	-0,72	1	-0,5	-0,5	0,61	0,23	-0,53	0,55	0,51
X3	0,33	0,33	0,77	0,63	0,63	0,58	0,5	0,88	-0,5	1	1	0,21	0,58	-0,18	0,21	0,2
X6	0,33	0,33	0,77	0,63	0,63	0,58	0,5	0,88	-0,5	1	1	0,21	0,58	-0,18	0,21	0,2
X10	-0,71	-0,71	-0,29	-0,42	-0,42	-0,21	-0,49	-0,1	0,61	0,21	0,21	1	0,79	-0,82	0,98	0,96
X11	-0,58	-0,58	0	-0,2	-0,2	0	-0,35	0,2	0,23	0,58	0,58	0,79	1	-0,87	0,77	0,76
X12	0,83	0,83	0,37	0,52	0,52	0,28	0,61	0,2	-0,53	-0,18	-0,18	-0,82	-0,87	1	-0,8	-0,79
X14	-0,68	-0,68	-0,22	-0,36	-0,36	-0,08	-0,45	-0,05	0,55	0,21	0,21	0,98	0,77	-0,8	1	0,99
X15	-0,67	-0,67	-0,19	-0,33	-0,33	-0,02	-0,44	-0,02	0,51	0,2	0,2	0,96	0,76	-0,79	0,99	1

Tabela 5.3

Apesar das novas variáveis continuam a poder formar-se três blocos de variáveis: analíticas, máximo e funções das dimensões.

Embora exista uma ligeira diminuição do primeiro valor próprio, a informação explicada pelas duas primeiras componentes é de 93%, o que significa que continuam a ser suficientes para traduzir a informação dos dados (Tabela 5.4).

Portanto, as variáveis introduzidas não contribuem com nova informação, o que aliás era manifesto pelas fortes correlações com as variáveis originais.

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
9,5573	59,73	59,73
4,9170	30,73	90,46
0,7667	4,79	95,26
0,4505	2,82	98,07
0,2133	1,33	99,41
0,0642	0,40	99,81
0,0187	0,12	99,92
0,0090	0,06	99,98
0,0024	0,01	100,00
0,0006	0,00	100,00
0,0002	0,00	100,00
0,0000	0,00	100,00
0,0000	0,00	100,00
0,0000	0,00	100,00
0,0000	0,00	100,00
0,0000	0,00	100,00
0,0000	0,00	100,00

Tabela 5.4

No entanto, apesar de não introduzirem alterações significativas, permitem uma mais clara definição dos grupos de paralelepípedos.

Observe-se a representação dos indivíduos no primeiro plano principal:

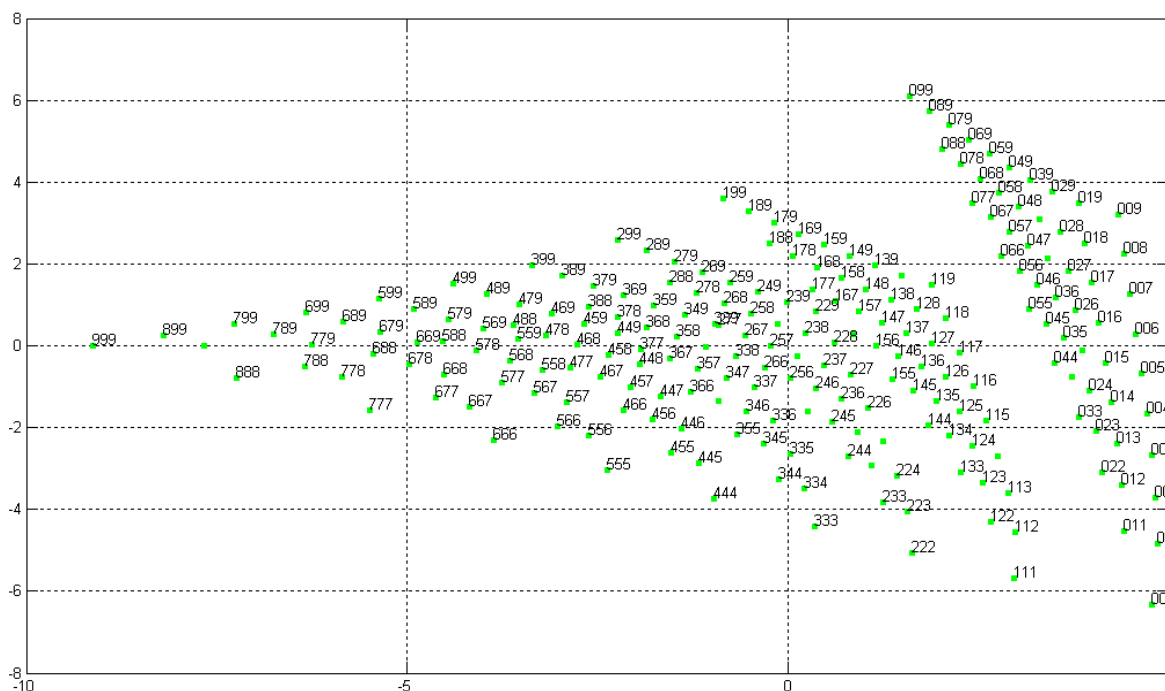


Fig. 5.1 - Indivíduos representados no plano principal (1,2)

Embora exista uma inversão em termos de organização nos dois eixos, a estrutura triangular dos indivíduos permanece, bem como a interpretação das respectivas arestas. Podem, contudo, notar-se grupos de paralelepípedos indistintos na abordagem inicial. Ou talvez, e mais correctamente, desapercebidos inicialmente.

O grupo mais à direita parece ser constituído por todos os paralelepípedos com dimensão mínima igual a 1. O segundo grupo é formado por todos os paralelepípedos de dimensão mínima 2 e assim sucessivamente. Estes grupos diluem-se com o aumentar da primeira dimensão, uma vez que vão sendo cada vez mais reduzidos. Neste sentido, as linhas equi-X1, abaixo apresentadas, são esclarecedoras:

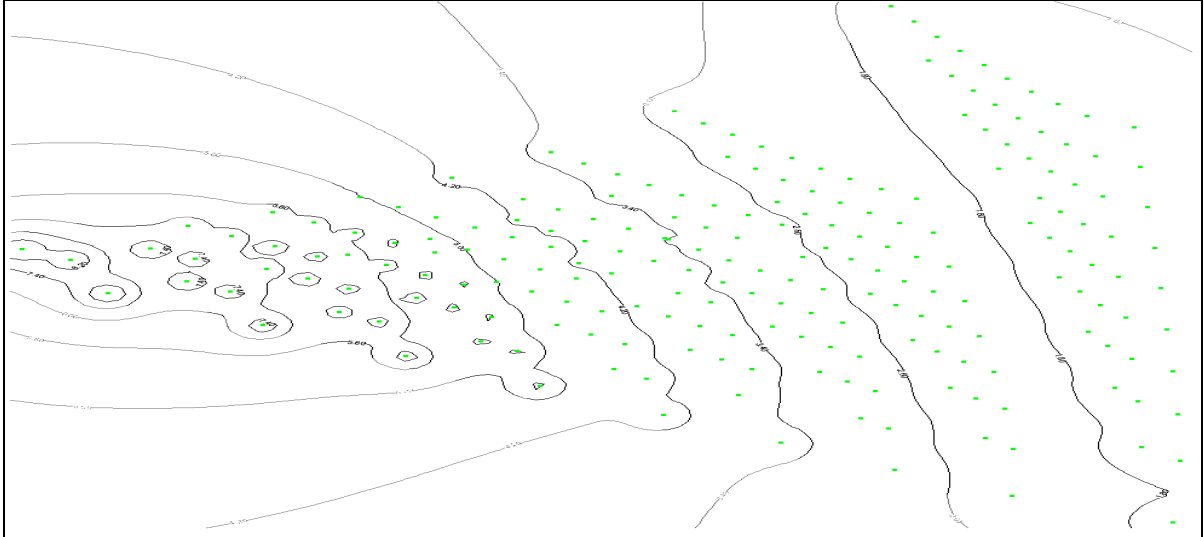


Fig. 5.2 - Linhas equi-X1 (12)

A mais evidente distinção parece ser causada pelas novas variáveis X_{13} e X_{16} , tendo em conta que se encontram muito correlacionadas com a primeira componente e as restantes correlações permanecem inalteradas (a menos de ligeiras diferenças). De facto observando o círculo e comparando com a situação inicial têm-se que, embora o sentido da correlação se altere, a organização é idêntica:

	Y1	Y2
X1	-0,9520	-0,2183
X2	-0,7670	0,4007
X3	-0,5180	0,7937
X4	-0,9426	0,0658
X5	-0,9569	0,2264
X6	-0,5180	0,7937
X7	-0,9520	-0,2183
X8	-0,9124	0,4021
X9	-0,9569	0,2264
X10	0,6296	0,7176
X11	0,3759	0,8764
X12	-0,6865	-0,6430
X13	-0,8153	0,5751
X14	0,5822	0,7411
X15	0,5565	0,7453
X16	0,8633	-0,0277

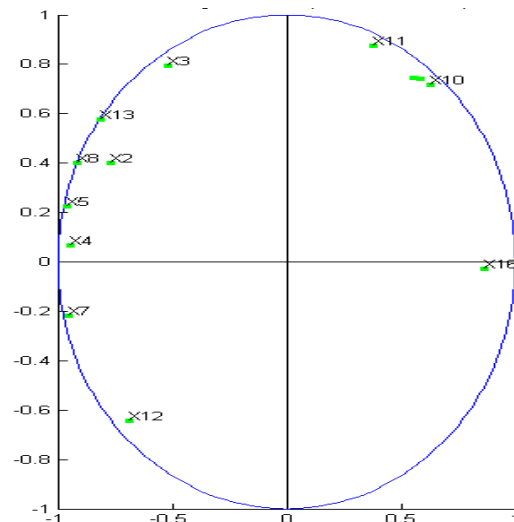


Fig. 5.3 - Círculo de Correlações (12)

Apresentam-se as linhas equi-volume que se comprova serem análogas às obtidas para as variáveis originais:

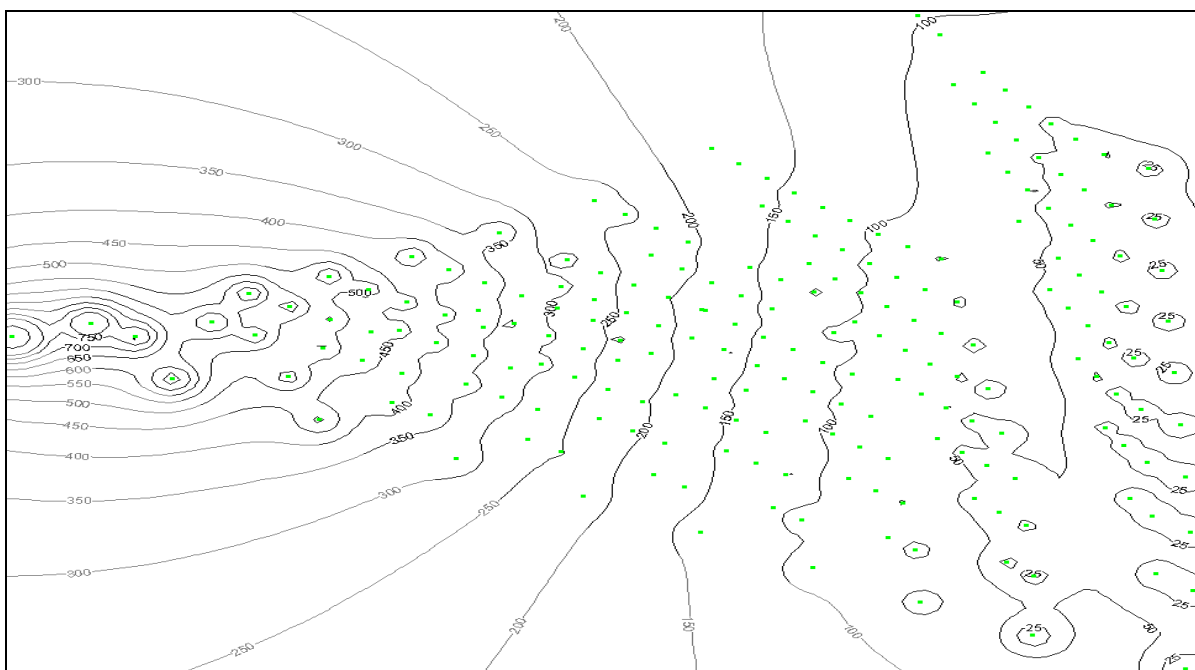


Fig. 5.4 - Linhas equi-volume (12)

A terceira componente tem agora uma interpretação diferente, já que as variáveis com maior importância para a sua formação são as novas variáveis X_{14} e X_{15} .

	Y3
X1	-0,0451
X2	-0,2854
X3	0,2825
X4	-0,1930
X5	-0,1451
X6	0,2825
X7	-0,0451
X8	-0,0357
X9	-0,1451
X10	-0,2041
X11	0,2837
X12	-0,1779
X13	-0,0022
X14	-0,3109
X15	-0,3402
X16	-0,2734

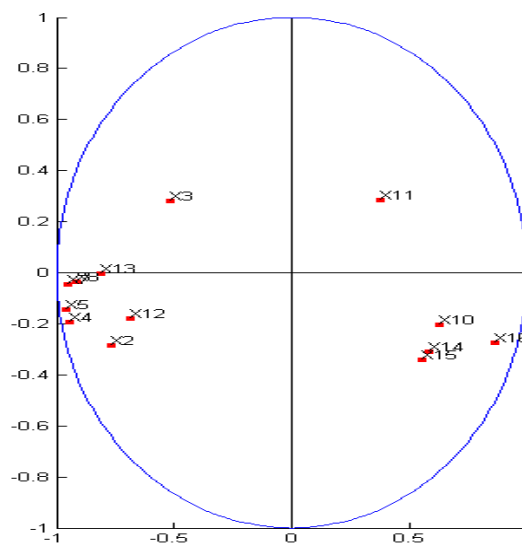


Fig. 5.5 - Círculo de Correlações (13)

Esta componente tem pouco peso em termos de inércia explicada (5% da inércia total).

Parece colocar em oposição paralelepípedos com duas dimensões iguais e grandes a paralelepípedos com duas dimensões iguais e pequenas.

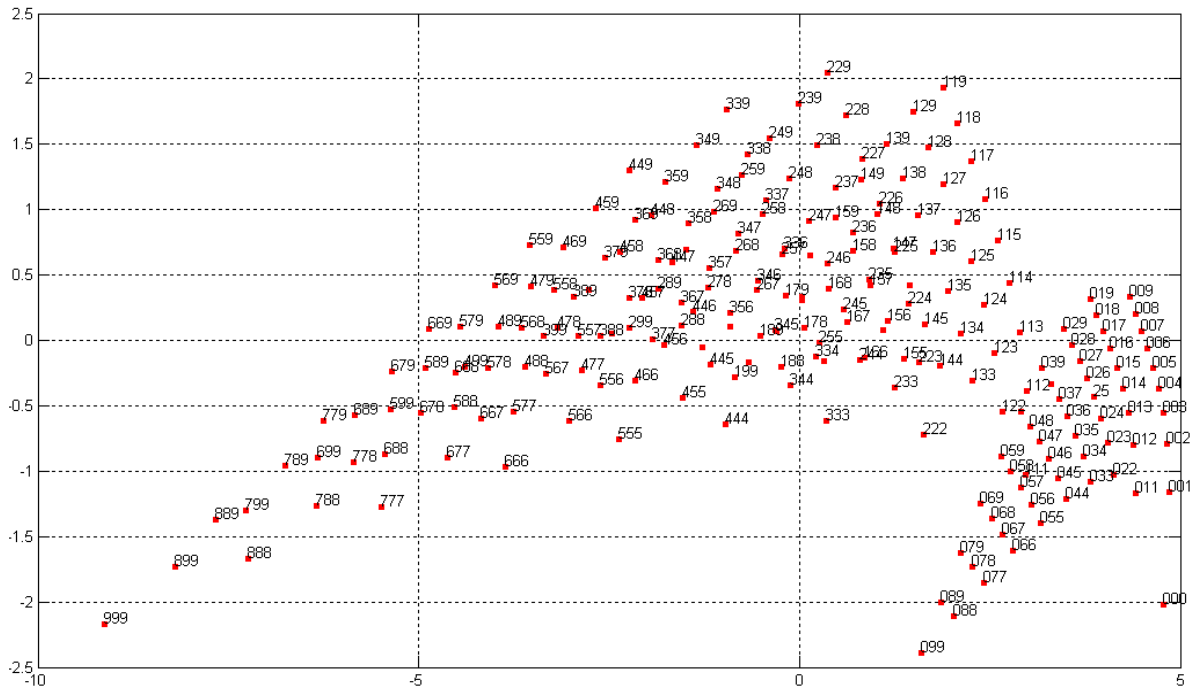


Fig. 5.6 - Indivíduos representados no plano principal (1,3)

Portanto, as variáveis introduzidas apenas trazem alterações significativas em termos do 3º eixo, que tem, no entanto, pouca importância para o total da inércia explicada.

Assim, as novas variáveis não parecem introduzir nova informação relevante. As variáveis inicialmente consideradas, embora de construção mais simples, são suficientes para captar o essencial da informação. Inclusivamente, as novas variáveis dificultam a distinção entre tábuas e palitos que anteriormente se encontravam em oposição no 3º eixo.

Permitem, contudo, nas duas primeiras componentes (nomeadamente, na primeira) notar grupos de paralelepípedos mais ‘disfarçados’ para as variáveis originais.

Além disso, mantém-se a preponderância das variáveis do maior bloco da matriz de correlações para a formação da primeira componente principal.

Capítulo VI

Matrizes Organizadas em Blocos

Consideremos a matriz de correlações $X = [n+m, n+m]$:

$$X = \begin{bmatrix} \sum_{11} & \sum_{12} \\ \sum_{21} & \sum_{22} \end{bmatrix}.$$

onde

$$\sum_{11} = \begin{bmatrix} 1 & \cdot & \cdot & 1-\varepsilon_1 \\ 1-\varepsilon_1 & \cdot & \cdot & 1 \end{bmatrix}, \quad \sum_{22} = \begin{bmatrix} 1 & \cdot & \cdot & 1-\varepsilon_2 \\ 1-\varepsilon_2 & \cdot & \cdot & 1 \end{bmatrix} \text{ e } \sum_{12} = \sum_{21}^T = \begin{bmatrix} \delta & \cdot & \cdot & \delta \\ \delta & \cdot & \cdot & \delta \end{bmatrix}$$

Para simplificar tomou-se $\varepsilon_1 = \varepsilon_2$.

Consideram-se então matrizes com dois blocos internamente muito correlacionados e pouco correlacionados entre si:

$$X = \begin{bmatrix} \begin{matrix} \mathbf{n} \\ \text{grandes} \end{matrix} & \begin{matrix} \text{pequenos} \end{matrix} \\ \begin{matrix} \text{pequenos} \end{matrix} & \begin{matrix} \text{grandes} \\ \mathbf{m} \end{matrix} \end{bmatrix}$$

Para matrizes de dimensão 12 testaram-se todos os tamanhos possíveis de sub-blocos: $n = 1 \dots 11$ e obteve-se o seguinte enunciado (não provado analiticamente):

Conjectura:

Os valores próprios de uma matriz $X = [n+m, n+m]$ nas condições anteriores são tais que $\lambda_3 = \lambda_4 = \dots = \lambda_{m+n} = \varepsilon$. Assim, como $\lambda_1 + \lambda_2 + (n+m-2)\varepsilon = n+m$ existe uma relação linear entre os dois maiores valores próprios: $\lambda_1 + \lambda_2 = (n+m)(1-\varepsilon) + 2\varepsilon$

Note-se que este resultado é independente da dimensão e do valor da menor correlação, δ .

Seguem-se alguns dos testes efectuados para diferentes valores de ε e δ .

Tomando $\varepsilon = 0.1$ e $\delta = 0.2$, obtiveram-se, de acordo com o valor de n , os seguintes valores próprios:

n	1	2	3	4	5	6	7	8	9	10	11
$\lambda_3 = \lambda_4 = \dots = \lambda_{12}$	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
λ_2	0,95137	1,7906	2,6069	3,3740	4,0134	4,3000	4,0134	3,3740	2,6069	1,7906	0,95137
λ_1	10,0490	9,2094	8,3931	7,6260	6,9866	6,7000	6,9866	7,6260	8,3931	9,2094	10,0490
$\lambda_1 + \lambda_2$	11,0004	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0004

Tabela 6.1

Considerando $\varepsilon = 0.1$ e $\delta = 0.3$, tem-se

n	1	2	3	4	5	6	7	8	9	10	11
$\lambda_3 = \lambda_4 = \dots = \lambda_{12}$	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
λ_2	0,89131	1,6581	2,3823	3,0261	3,5100	3,700	3,5100	3,0261	2,3823	1,6581	0,89131
λ_1	10,109	9,3419	8,6177	7,9739	7,4900	7,300	7,4900	7,9739	8,6177	9,3419	10,109
$\lambda_1 + \lambda_2$	11,0003	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0000	11,0003

Tabela 6.2

Ou seja, quer para $\delta=0,2$, quer para $\delta=0,3$ obteve-se $\lambda_1 + \lambda_2 = 11$, o que está de acordo com a conjectura enunciada:

$$\lambda_1 + \lambda_2 = 12 \times 0,9 + 2 \times 0,1 = 11 \Leftrightarrow \lambda_1 = -\lambda_2 + 11$$

A relação linear entre os dois maiores valores próprios (com $\varepsilon = 0,1$ e $\delta = 0,3$) fica explícita no gráfico seguinte:

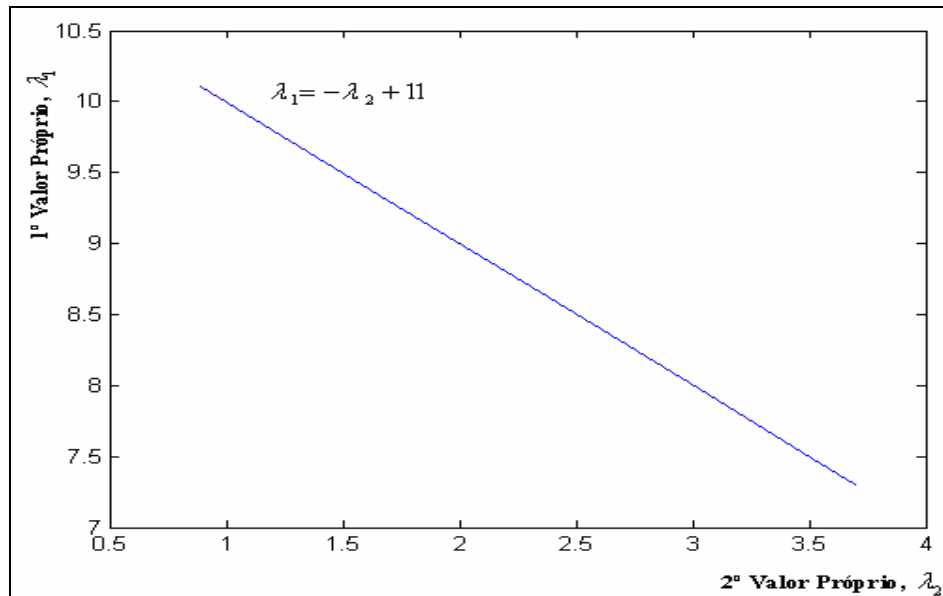


Fig. 6.1- Representação λ_1 vs λ_2 correspondente aos valores da tabela 6.2

Tomando agora $\varepsilon = 0,2$, tem-se $\lambda_1 + \lambda_2 = 12 \times 0,8 + 2 \times 0,2 = 10 \Leftrightarrow \lambda_2 = -\lambda_1 + 10$.

Novamente, para $\delta = 0,2$ e $0,3$ é verificada a conjectura, como se comprova pelos valores das tabelas 6.3 e 6.4:

n	1	2	3	4	5	6	7	8	9	10	11
$\lambda_3 = \lambda_4 = \dots = \lambda_{12}$	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2
λ_2	0,94537	1,6774	2,3847	3,0404	3,5717	3,8000	3,5717	3,0404	2,3847	1,6774	0,94537
λ_1	9,0546	8,3226	7,6153	6,9596	6,4283	6,2000	6,4283	6,9596	7,6153	8,3226	9,0546

$\lambda_1 + \lambda_2$	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000
-------------------------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------

Tabela 6.3

n	1	2	3	4	5	6	7	8	9	10	11
$\lambda_3 = \lambda_4 = \dots = \lambda_{12}$	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2
λ_2	0,87811	1,5301	2,1382	2,6676	3,0532	3,2000	3,0532	2,6676	2,1382	1,5301	0,87811
λ_1	9,1219	8,4699	7,8618	7,3324	6,9468	6,8000	6,9468	7,3324	7,8618	8,4699	9,1219

$\lambda_1 + \lambda_2$	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000	10,0000
-------------------------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------

Tabela 6.4

O gráfico λ_1 vs λ_2 , referente aos valores da Tabela 6.3, revela a linearidade da relação entre os dois valores próprios:

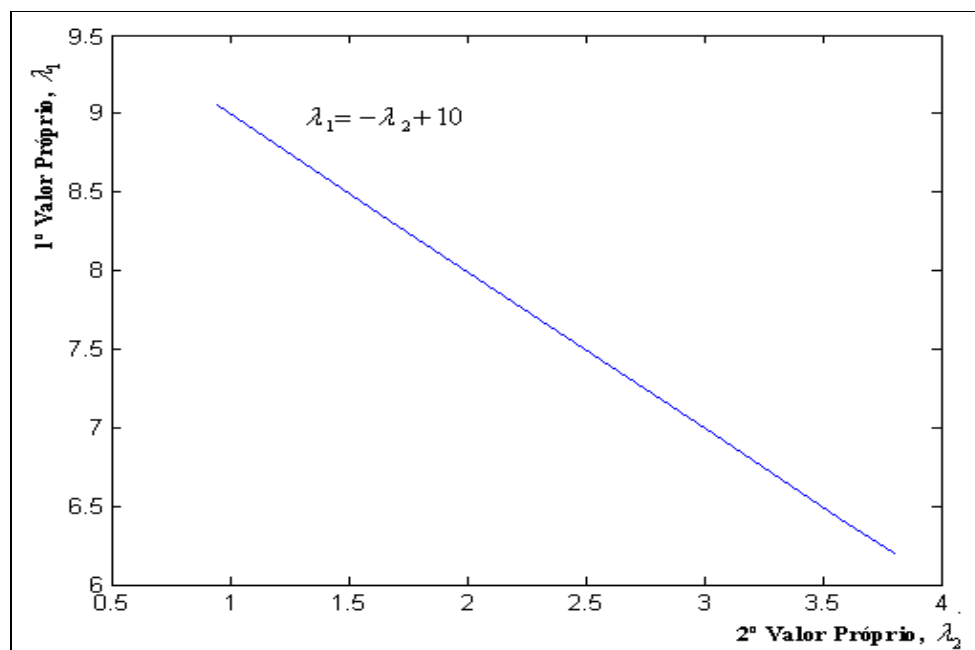


Fig. 6.2 - Representação λ_1 vs λ_2 correspondente aos valores da tabela 6.3

Para $\varepsilon = 0,3$ observa-se de novo a relação linear entre λ_1 e λ_2 :

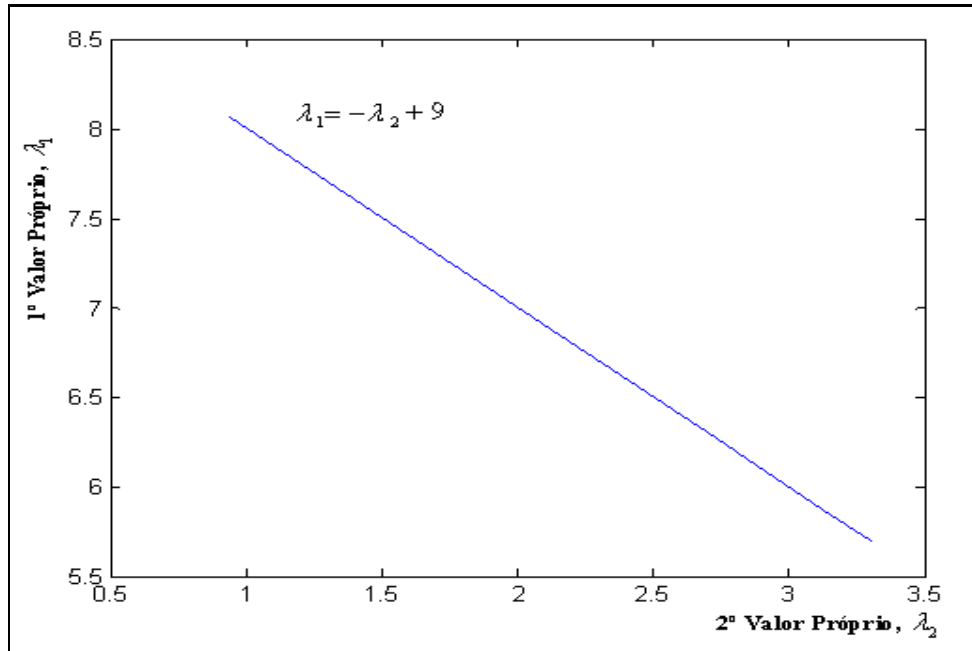


Fig. 6.3- Representação λ_1 vs λ_2 correspondente aos valores da tabela 6.5

Neste caso, $\lambda_1 + \lambda_2 = 12 \times 0,7 + 2 \times 0,3 = 9 \Leftrightarrow \lambda_1 = -\lambda_2 + 9$:

n	1	2	3	4	5	6	7	8	9	10	11
$\lambda_3 = \lambda_4 = \dots = \lambda_{12}$	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3
λ_2	0,9377	1,5606	2,1569	2,7000	3,1252	3,3000	3,1252	2,7000	2,1569	1,5606	0,9377
λ_1	8,0623	7,4394	6,8431	6,3000	5,8748	5,7000	5,8748	6,3000	6,8431	7,4394	8,0623
$\lambda_1 + \lambda_2$	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000

Tabela 6.5

n	1	2	3	4	5	6	7	8	9	10	11
$\lambda_3 = \lambda_4 = \dots = \lambda_{12}$	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3
λ_2	0,86132	1,3952	1,8847	2,3000	2,5921	2,7000	2,5921	2,3000	1,8847	1,3952	0,86132
λ_1	8,1387	7,6048	7,1153	6,7000	6,4079	6,3000	6,4079	6,7000	7,1153	7,6048	8,1387
$\lambda_1 + \lambda_2$	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000	9,0000

Tabela 6.6

Para além das matrizes de dimensão 12 foram testadas matrizes de dimensão 100 e 200 e diferentes valores de δ .

Apresenta-se (Tabela 6.7) a título de exemplo o resultado obtido considerando uma matriz de dimensão 100 com $\varepsilon = 0,1$ e $\delta = 0,1$ (embora para o resultado em causa o valor de δ se torne irrelevante).

n	1	2	5	10	15	20	25	30	40	50	60 ...
$\lambda_3 = \lambda_4 = \dots = \lambda_{100}$	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
λ_2	0,9888	1,8773	4,5414	8,9752	13,3980	17,8050	22,1870	26,5260	34,8530	40,1000	34,8530
λ_1	89,2110	88,3230	85,6590	81,2250	76,8020	72,3950	68,0130	63,6740	55,3470	50,1000	55,3470
$\lambda_1 + \lambda_2$	90,1998	90,2003	90,2004	90,2000	90,2000	90,2000	90,2000	90,2000	90,2000	90,2000	90,2000

Tabela 6.7

Tem-se então $\lambda_1 + \lambda_2 = 100 \times 0,9 + 2 \times 0,1 = 90,2 \Leftrightarrow \lambda_2 = -\lambda_1 + 90,2$:

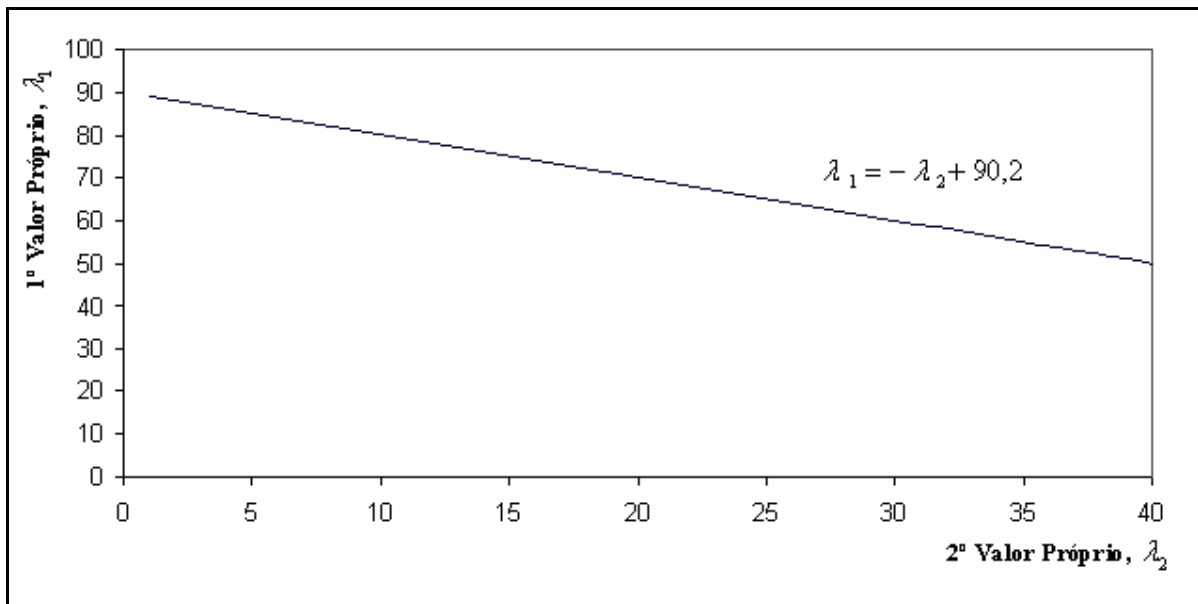


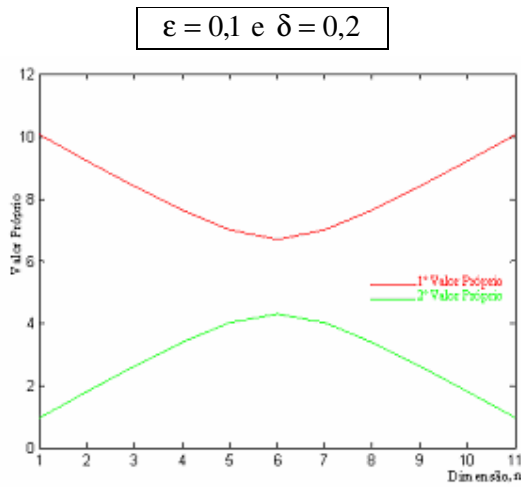
Fig. 6.4 - Representação λ_1 vs λ_2 correspondente aos valores da tabela 6.7

6.2 Variação dos Maiores Valores Próprios em Função da Dimensão dos Blocos, de ε e δ

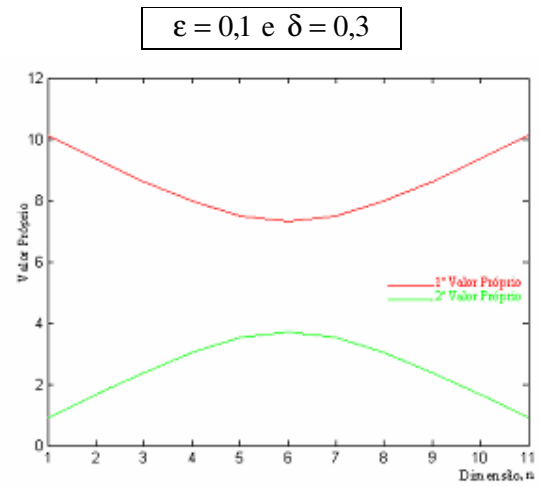
Para matrizes de correlação de dimensão 12 seguem-se os testes efectuados para o estudo da variação dos maiores valores próprios em função de n , ε e δ . Tomaram-se para ε e δ os valores anteriormente considerados e $n = 1..11$.

Para cada um dos casos é calculado o polinómio de 2º grau que melhor aproxima, no sentido dos mínimos quadrados, a curva obtida.

Seguem-se os resultados obtidos em cada caso. Os polinómios (I) e (II) representam, respectivamente, a variação do maior e menor valor próprio em função de n .

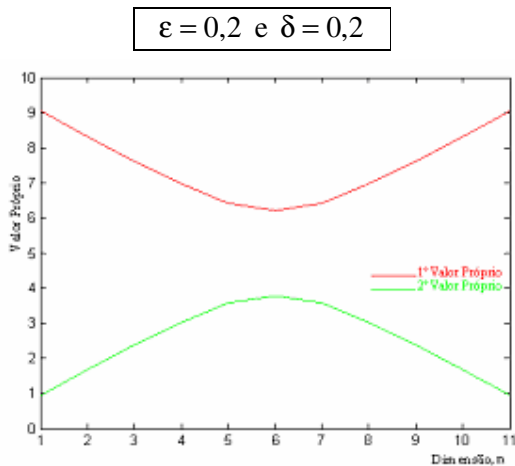


(I) $y = 0,1293n^2 - 1,5513n + 11,6545$
 (II) $y = -0,1293n^2 + 1,5513n - 0,6545$

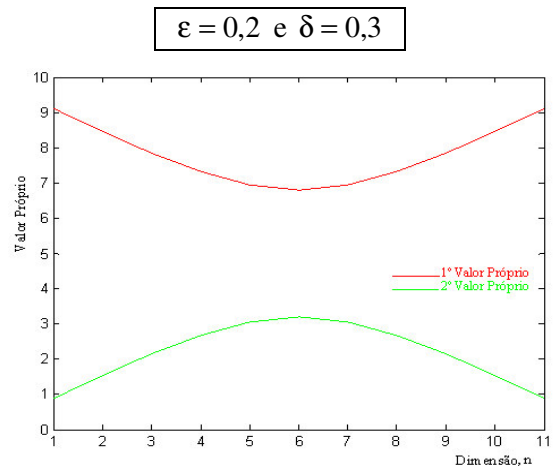


(I) $y = 0,1103n^2 - 1,3234n + 11,4459$
 (II) $y = -0,1103n^2 + 1,3234n - 0,4459$

Fig. 6.5 – Variação dos Maiores Valores Próprios em função da dimensão

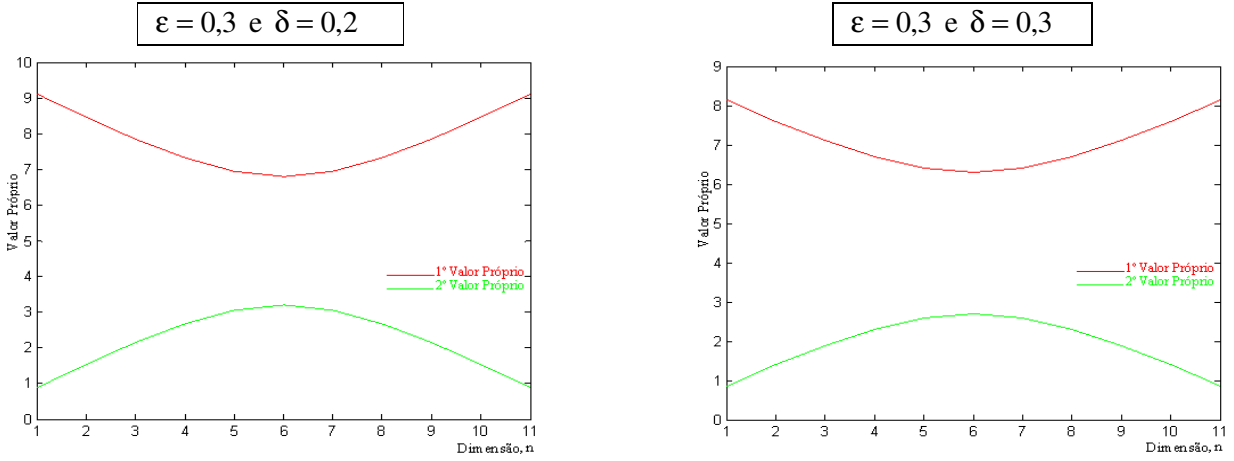


(I) $y = 0,1108n^2 - 1,3295n + 10,4224$
 (II) $y = -0,1108n^2 + 1,3295n - 0,4224$



(I) $y = 0,0915n^2 - 1,0985n + 10,2223$
 (II) $y = -0,0915n^2 + 1,0985n - 0,2223$

Fig. 6.66 – Variação dos Maiores Valores Próprios em função da dimensão



$$(I) \quad y = 0,0922n^2 - 1,1064n + 9,1917$$

$$(II) \quad y = -0,0922n^2 + 1,1064n - 0,1917$$

$$(I) \quad y = 0,0728n^2 - 0,8734n + 9,0044$$

$$(II) \quad y = -0,0728n^2 + 0,8734n - 0,0044$$

Fig. 6.7 - Variação dos Maiores Valores Próprios em função da dimensão

Procurou-se, a partir dos resultados obtidos através dos mínimos quadrados, obter um polinómio função de n , ε e δ . Notar que para além dos testes anteriormente apresentados foram também testadas matrizes de dimensão 100 e 200.

$$\text{Sejam } A = \left(\frac{(1-\varepsilon)-\delta}{\frac{n+m}{2}} + \frac{\delta}{\frac{10(n+m)}{2}} \right), \text{ e } B = (n+m)(1-\varepsilon) + 2\varepsilon.$$

Concluiu-se que

$$\lambda_1(n; \varepsilon, \delta) = A \times \left(n - \frac{n+m}{2} \right)^2 - A \times \left(\frac{n+m}{2} \right)^2 + B + E(n, \varepsilon, \delta) + E2$$

$$\lambda_2(n; m, \delta) = -A \times \left(n - \frac{n+m}{2} \right)^2 + A \times \left(\frac{n+m}{2} \right)^2 - E(n, m, \varepsilon, \delta) - E2$$

$E(n, \varepsilon, \delta)$ representa um erro função da dimensão, ε e δ , cujo valor é, em todos os casos conhecido. E2 um erro de aproximação, pois verificou-se que para dimensões superiores (em particular, 100 e 200), um polinómio de grau 4 constitui uma melhor aproximação que um polinómio de 2º grau. Exemplo desta situação será apresentado adiante neste capítulo.

Seguem os resultados para o primeiro valor próprio calculados através do polinómio obtido e os valores esperados (Tabela 6.8).

n	$\varepsilon = 0,1$ e $\delta = 0,2$		$\varepsilon = 0,1$ e $\delta = 0,3$		$\varepsilon = 0,2$ e $\delta = 0,2$		$\varepsilon = 0,2$ e $\delta = 0,3$		$\varepsilon = 0,3$ e $\delta = 0,2$		$\varepsilon = 0,3$ e $\delta = 0,3$	
	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado
1	10,3345	10,049	10,2909	10,109	9,2877	9,0546	9,2506	9,1219	8,2284	8,0623	8,2161	8,1387
2	9,2545	9,2094	9,3459	9,3419	8,3577	8,3226	8,4556	8,4699	7,4484	7,4394	7,5711	7,6048
3	8,4145	8,3931	8,6109	8,6177	7,6344	7,6153	7,8373	7,8618	6,8417	6,8431	7,0694	7,1153
4	7,8145	7,6260	8,0859	7,9739	7,1177	6,9596	7,3956	7,3324	6,4084	6,3000	6,7111	6,7000
5	7,4545	6,9866	7,7709	7,4900	6,8077	6,4283	7,1306	6,9468	6,1484	5,8748	6,4961	6,4079
6	7,3345	6,7000	7,6659	7,3000	6,7044	6,2000	7,0423	6,8000	6,0617	5,7000	6,4244	6,3000
7	7,4545	6,9866	7,7709	7,4900	6,8077	6,4283	7,1306	6,9468	6,1484	5,8748	6,4961	6,4079
8	7,8145	7,6260	8,0859	7,9739	7,1177	6,9596	7,3956	7,3324	6,4084	6,3000	6,7111	6,7000
9	8,4145	8,3931	8,6109	8,6177	7,6344	7,6153	7,8373	7,8618	6,8417	6,8431	7,0694	7,1153
10	9,2545	9,2094	9,3459	9,3419	8,3577	8,3226	8,4556	8,4699	7,4484	7,4394	7,5711	7,6048
11	10,3345	10,0490	10,2909	10,1090	9,2877	9,0546	9,2506	9,1219	8,2284	8,0623	8,2161	8,1387

Tabela 6.8

Verifica-se que os resultados são bastante satisfatórios, embora para blocos de dimensão intermédia o erro cometido seja maior. Apresenta-se abaixo o gráfico onde são comparados os valores obtidos e valores esperados para matrizes de dimensão 12 (com $\varepsilon = 0,1$ e $\delta = 0,2$) e onde, de facto, se verifica, que para dimensões intermédias o erro cometido é maior.

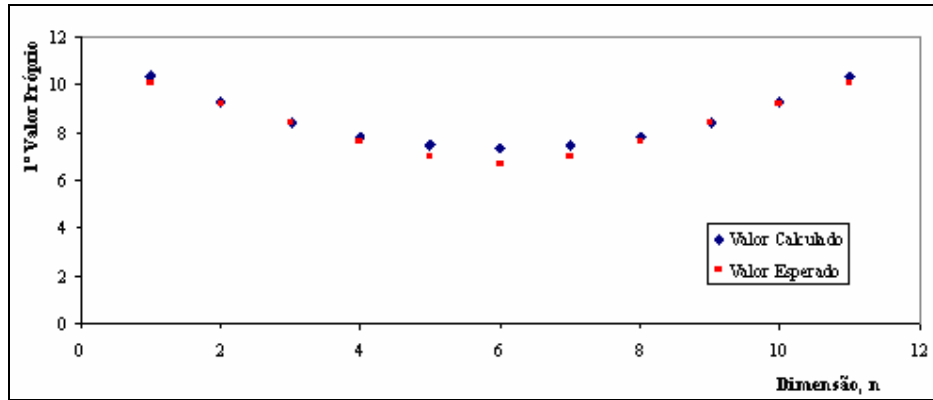


Fig. 6.8 – Comparação do 1º Valor Próprio Calculado e Esperado

Para matrizes de dimensão 100 o erro cometido na aproximação aumenta ligeiramente como se pode verificar através dos resultados obtidos:

n	$\varepsilon = 0,1$ e $\delta = 0,2$		$\varepsilon = 0,1$ e $\delta = 0,3$		$\varepsilon = 0,2$ e $\delta = 0,2$		$\varepsilon = 0,2$ e $\delta = 0,3$		$\varepsilon = 0,3$ e $\delta = 0,2$		$\varepsilon = 0,3$ e $\delta = 0,3$	
	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado
1	92,7504	89,2450	91,8176	89,3010	82,3390	79,4500	81,4476	79,5130	71,9343	69,6580	71,1084	69,7300
2	91,3536	88,3910	90,5954	88,5040	81,1362	78,7020	80,4194	78,8290	70,9255	69,0160	70,2742	69,1610
3	89,9856	87,5370	89,3984	87,7080	79,9582	77,9540	79,4124	78,1470	69,9375	68,3760	69,4572	68,5960
10	81,216	81,5970	81,7250	82,2080	72,4066	72,7580	72,9570	73,4420	63,6039	63,9360	64,2198	64,7110
11	80,0784	80,7530	80,7296	81,4330	71,4270	72,0210	72,1196	72,7810	51,1239	52,0330	53,8998	54,9210
20	71,1360	73,2600	72,9050	74,6470	63,7266	65,4980	65,5370	67,0330	49,3039	49,4860	52,3948	53,0480
30	63,9360	65,2990	66,6050	67,7500	57,5266	58,6390	60,2370	61,2950	48,0039	47,3420	51,3198	51,5790
35	61,4160	61,6300	64,4000	64,7720	55,3566	55,5300	58,3820	58,8750	46,9639	45,30	50,4598	50,3000
40	59,6160	58,4040	62,8250	62,3340	53,8066	52,8490	57,0570	56,9330	71,9343	69,6580	71,1084	69,7300
50	58,1760	55,1000	61,5650	60,1000	52,5666	50,200	55,9970	55,2000	70,9255	69,0160	70,2742	69,1610

Tabela 6.9

De facto, o gráfico comparativo dos valores calculados e esperados (em particular, para $\varepsilon = 0,1$ e $\delta = 0,2$) revela ligeiro afastamento nos extremos, novo afastamento após aproximação e um maior afastamento entre os valores nas dimensões intermédias.

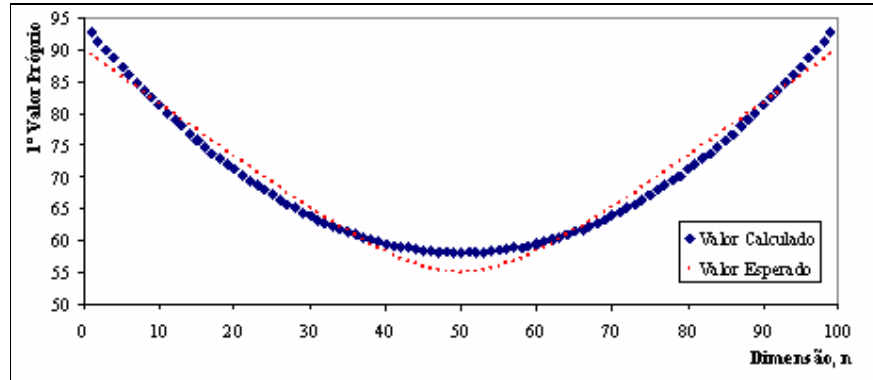


Fig. 6.9 – Comparação do 1º Valor Próprio Calculado e Esperado

Para matrizes de dimensão 200 o resultado é análogo ao obtido para matrizes de dimensão 100:

n	$\varepsilon = 0,1$ e $\delta = 0,2$		$\varepsilon = 0,1$ e $\delta = 0,3$		$\varepsilon = 0,2$ e $\delta = 0,2$		$\varepsilon = 0,2$ e $\delta = 0,3$		$\varepsilon = 0,3$ e $\delta = 0,2$		$\varepsilon = 0,3$ e $\delta = 0,3$	
	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado	Calculado	Esperado
1	186,5772	179,2400	184,5992	179,3000	165,5014	159,4500	159,3453	159,5100	144,4164	139,6600	141,1041	139,7300
2	185,1588	178,3900	183,3581	178,5000	164,2800	158,7000	158,3012	158,8300	143,3920	139,0200	140,2570	139,1600
3	183,7548	177,5400	182,1296	177,7000	163,0710	157,9500	157,2677	158,1400	142,378	138,3700	139,4185	138,5900
19	163,2492	163,9400	164,1872	165,0900	145,4134	146,0500	142,1733	147,3500	127,5684	128,2000	127,1721	129,6700
20	162,0900	163,0900	163,1729	164,3200	144,4152	145,3200	141,3200	146,6800	126,7312	127,5700	126,4798	129,1200
50	134,0100	138,3200	138,6029	142,0600	120,2352	123,7900	120,6500	127,9000	106,4512	109,3500	109,7098	113,8900
70	122,4900	123,1600	128,5229	129,4400	110,3152	110,8600	112,1700	117,5500	98,1312	98,6730	102,8298	105,8000
71	122,0652	122,4700	128,1512	128,9000	109,9494	110,2800	111,8573	117,1100	97,8244	98,2010	102,5761	105,4600
80	118,8900	116,7100	125,3729	124,5700	107,2152	105,5000	109,5200	113,6700	95,5312	94,3830	100,6798	102,8600
100	116,0100	110,1000	122,8529	120,1000	104,7352	100,2000	107,4000	110,2000	93,4512	90,3000	98,9598	100,3000

Tabela 6.10

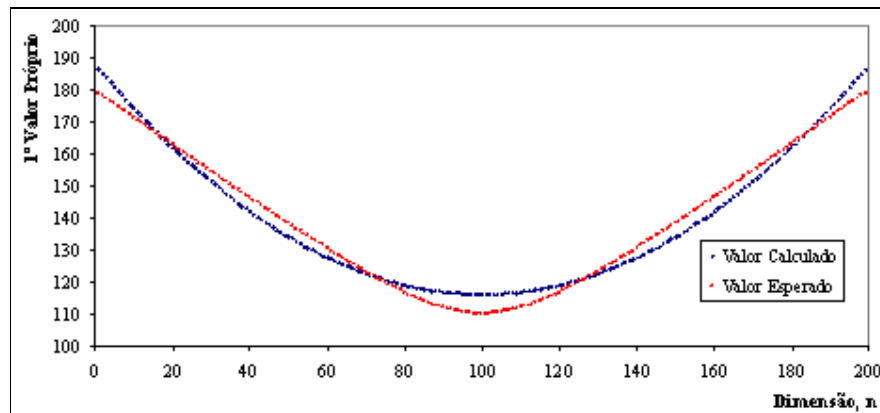


Fig. 6.10 - Comparação do 1º Valor Próprio Calculado e Esperado

Justifica-se o aumento do erro de aproximação do primeiro valor próprio de matrizes de maior dimensão pelo facto, já referido, um polinómio de grau 4, por exemplo, constituir uma melhor aproximação que um polinómio de 2º grau (no sentido dos mínimos quadrados):

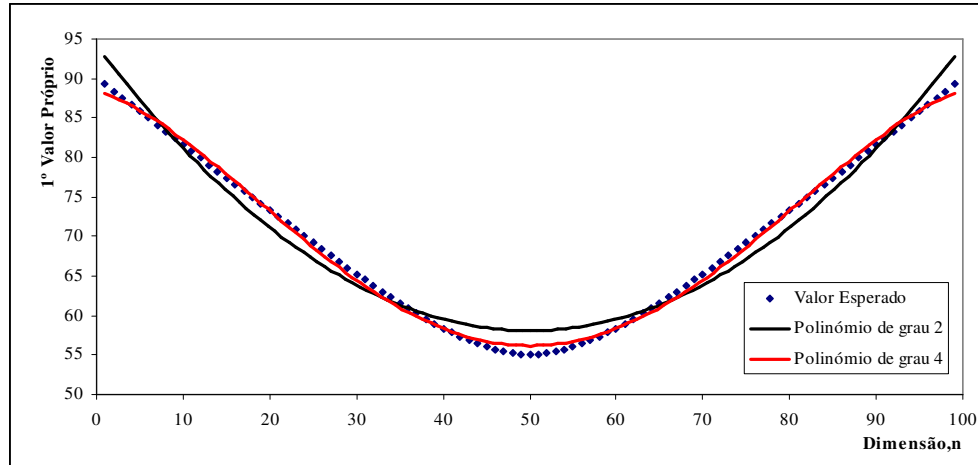


Fig. 6.11 - Comparação da aproximação feita por polinómio de grau 2 e grau 4 para matrizes de dimensão 100 com $\varepsilon = 0,1$ e $\delta = 0,2$

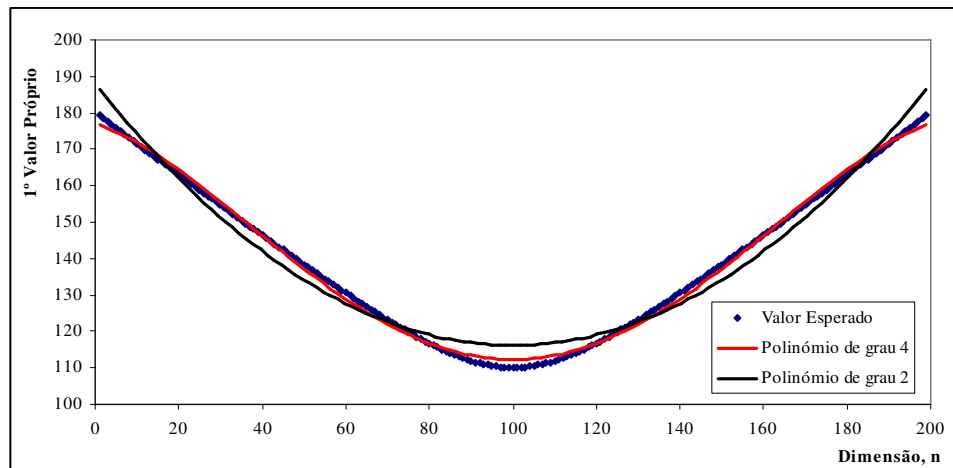


Fig. 6.12 - Comparação da aproximação feita por polinómio de grau 2 e grau 4 para matrizes de dimensão 200 com $\varepsilon = 0,1$ e $\delta = 0,2$

Considera-se no entanto que o polinómio obtido para o cálculo do 1º valor próprio, em função de n , ε e δ , representa uma boa aproximação do valor real.

6.3 Organização do Círculo de Correlações

Analisa-se de seguida o comportamento das variáveis no círculo de correlações, quando a matriz de correlações associada se organiza em dois blocos fortemente correlacionados.

Tomaram-se valores de $\varepsilon = 0,1; 0,2$ e $0,3$ e $\delta = 0,1; 0,2$.

Nas figuras 6.13, 6.14 e 6.15 são apresentados os círculos de correlações obtidos para os diferentes valores de ε e δ .

Em qualquer dos casos considerados verifica-se que no círculo se distinguem dois grupos de variáveis. Cada um destes grupos está associada a um dos blocos de grandes correlações.

Naturalmente que variando o valor de n também os grupos de variáveis se vão alterando. Nos círculos de correlações apresentados, os grupos I e II indicados correspondem, consoante o valor de n , aos seguintes grupos de variáveis (Tabela 6.11):

	Grupo I	Grupo II
$n = 1$	$\{X_1\}$	$\{X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 2$	$\{X_1, X_2\}$	$\{X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 3$	$\{X_1, X_2, X_3\}$	$\{X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 4$	$\{X_1, X_2, X_3, X_4\}$	$\{X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 5$	$\{X_1, X_2, X_3, X_4, X_5\}$	$\{X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 6$	$\{X_1, X_2, X_3, X_4, X_5, X_6\}$	$\{X_7, X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 7$	$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7\}$	$\{X_8, X_9, X_{10}, X_{11}, X_{12}\}$
$n = 8$	$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8\}$	$\{X_9, X_{10}, X_{11}, X_{12}\}$
$n = 9$	$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9\}$	$\{X_{10}, X_{11}, X_{12}\}$
$n = 10$	$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}\}$	$\{X_{11}, X_{12}\}$
$n = 11$	$\{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}\}$	$\{X_{12}\}$

Tabela 6.11

Pode observar-se que se encontram mais fortemente correlacionadas com a primeira componente as variáveis pertencentes ao bloco de maior dimensão. O bloco de menor dimensão correlaciona-se mais fortemente com a segunda componente. Daí que aumentando a dimensão de n o grupo I se desloquem na direcção do eixo dos xx e o grupo II no sentido do eixo dos yy .

Note-se que aumentando o valor de ε as variáveis se aproximam do interior do círculo. Por outro lado, para o mesmo valor de ε , aumentando o valor de δ diminui o ângulo entre as variáveis do primeiro grupo e o primeiro eixo e aumenta o ângulo entre este eixo e as variáveis do segundo grupo.

Também se observa que as variáveis pertencentes ao bloco superior esquerdo estão correlacionadas negativamente com o primeiro eixo, ao invés das variáveis pertencentes ao bloco inferior direito que se correlacionam com este eixo positivamente.

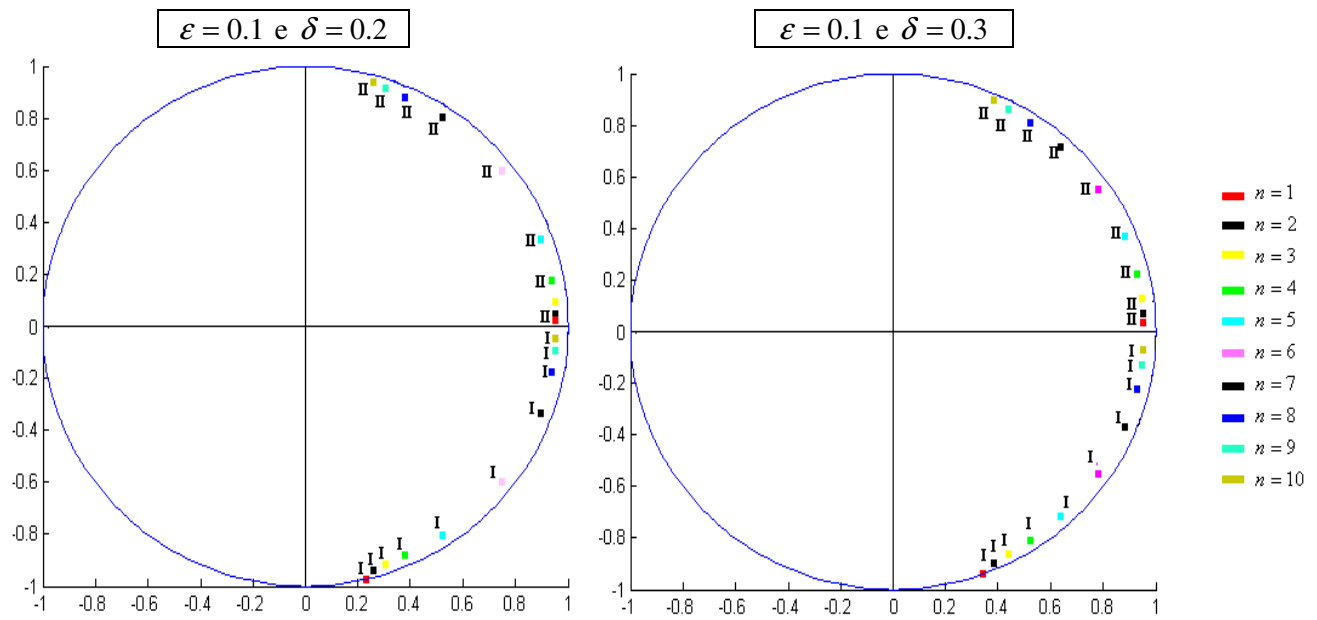


Fig. 6.13 – Círculo de Correlações (12)

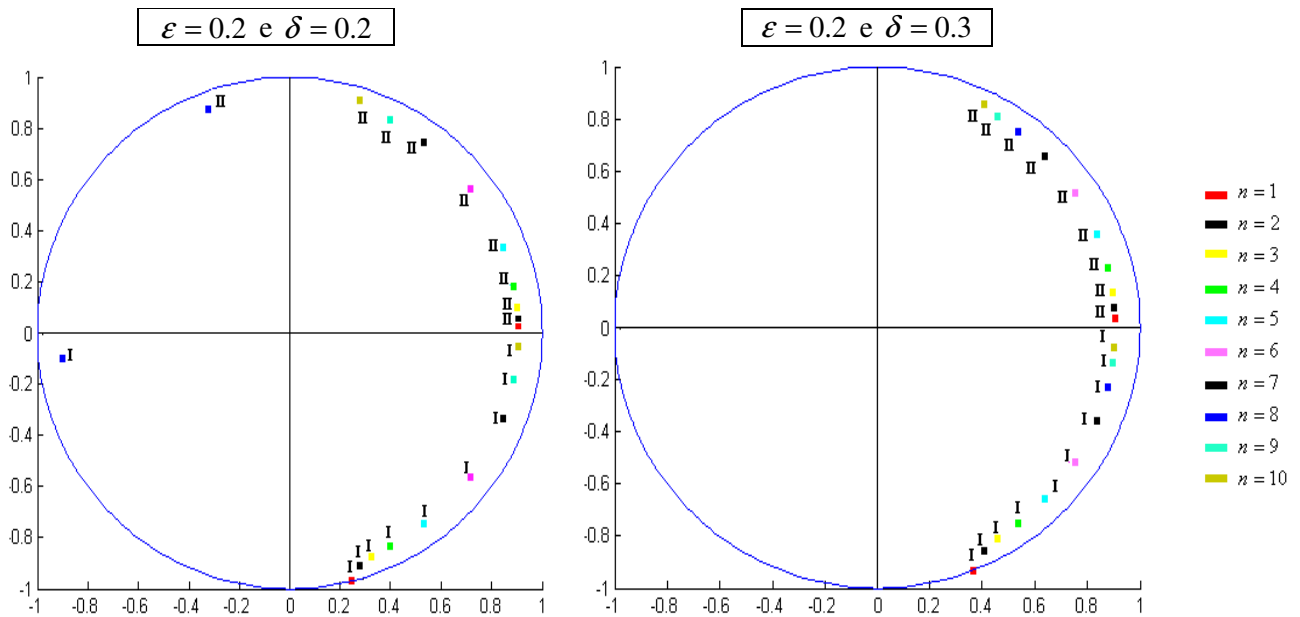


Fig. 6.13 - Círculo de Correlações (12)

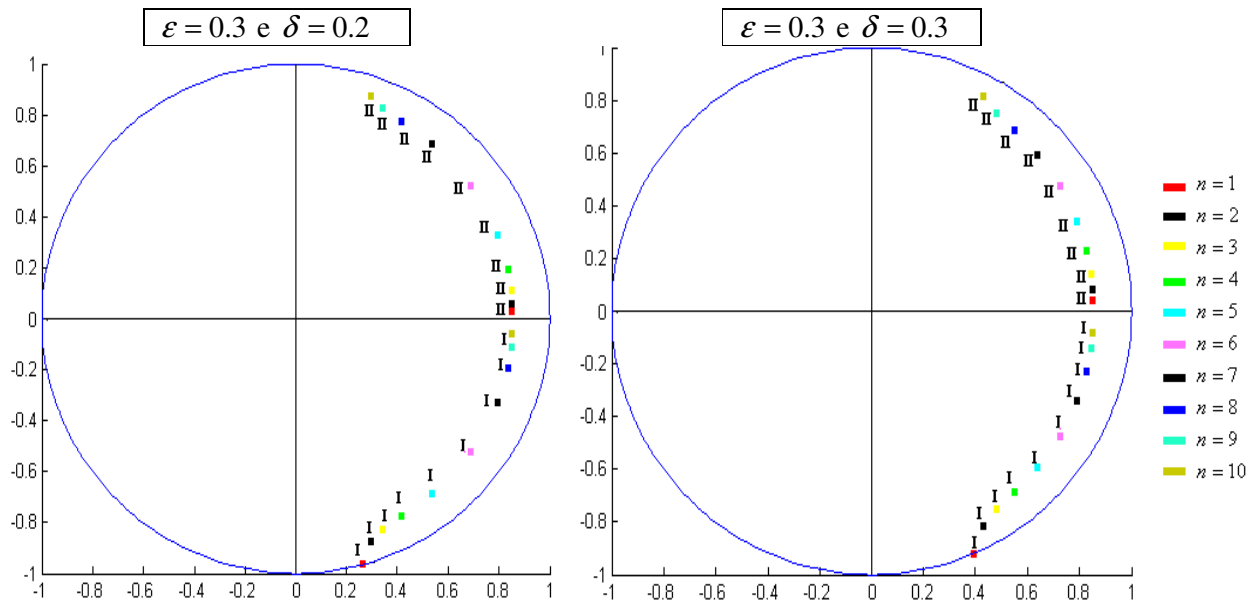
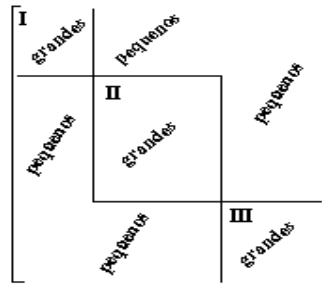


Fig. 6.15 - Círculo de Correlações (12)

Estudaram-se também matrizes organizadas em três blocos de fortes correlações:



Supondo primeiro que o número de variáveis de cada bloco é idêntico ao que se obteve para matriz de correlações apresentada no estudo dos 220 paralelepípedos, ou seja, bloco I com 7 variáveis, bloco II com 3 variáveis e bloco III com duas variáveis. Para $\delta = 0.2$ e tomando $\varepsilon = 0.1; 0.2$, obtiveram-se as seguintes representações

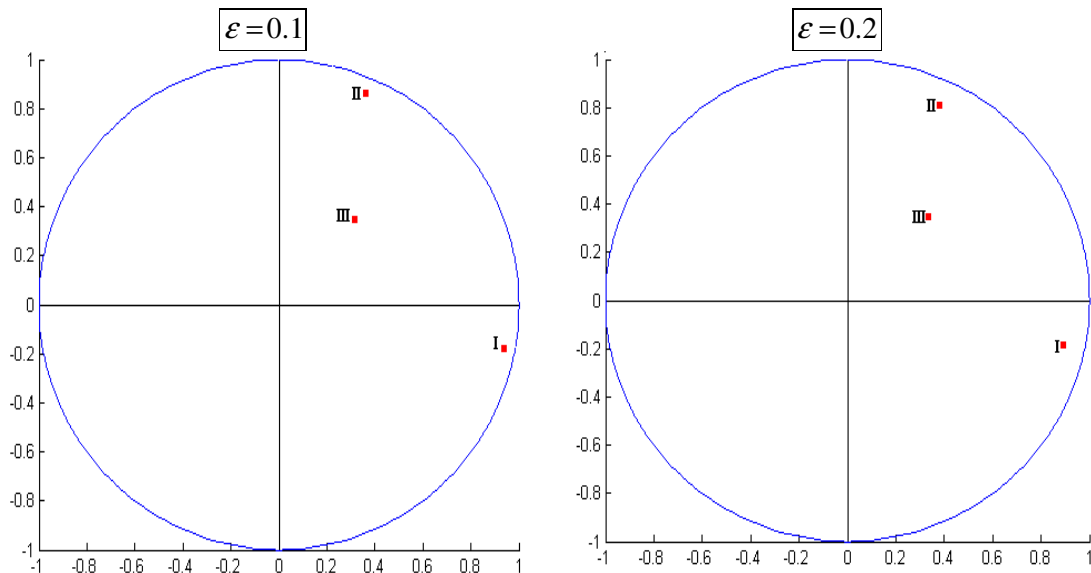


Fig. 6.16 - Círculo de Correlações (12)

Verifica-se que é o primeiro bloco, ou seja, o bloco de maior dimensão, que mais contribui para a formação da primeira componente. O bloco II é aquele que mais contribui para a formação do segundo eixo.

Observa-se igualmente que diminuindo o valor das correlações dos blocos os grupos tendem a dirigir-se para o interior do círculo.

Efectua-se novo teste, agora com as seguintes dimensões para os blocos: bloco I com 3 variáveis, bloco II com 7 e bloco III com 2:

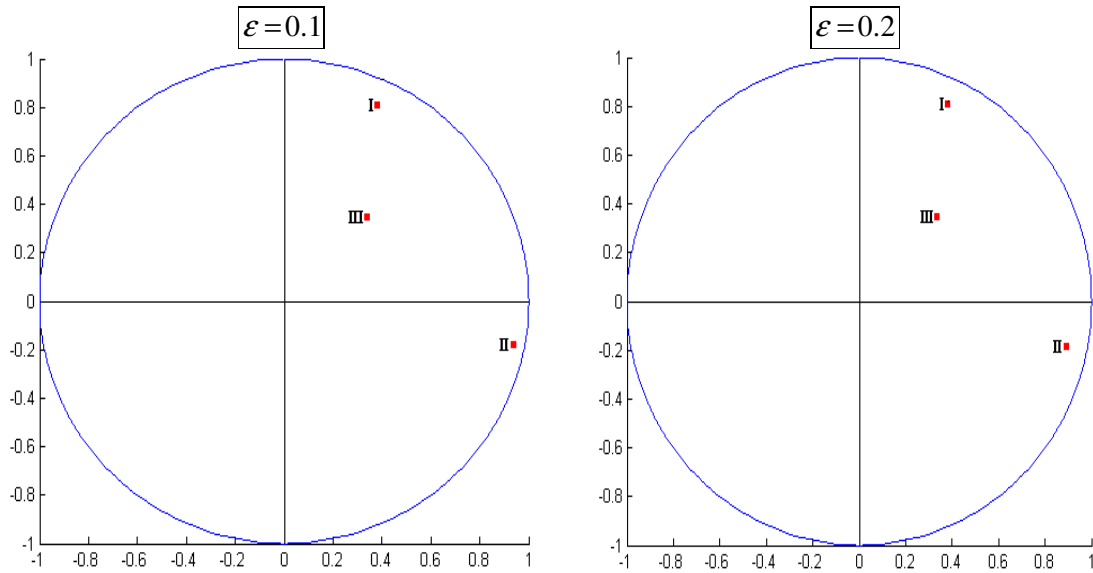


Fig. 6.17 -Círculo de Correlações (12)

As conclusões mantêm-se, isto é, o maior bloco de fortes correlações é o que mais preponderante para a formação do primeiro eixo, e o segundo maior é o mais influente na formação do segundo eixo.

Examina-se uma última situação: bloco III com 6 variáveis e os blocos I e II com 3 variáveis.

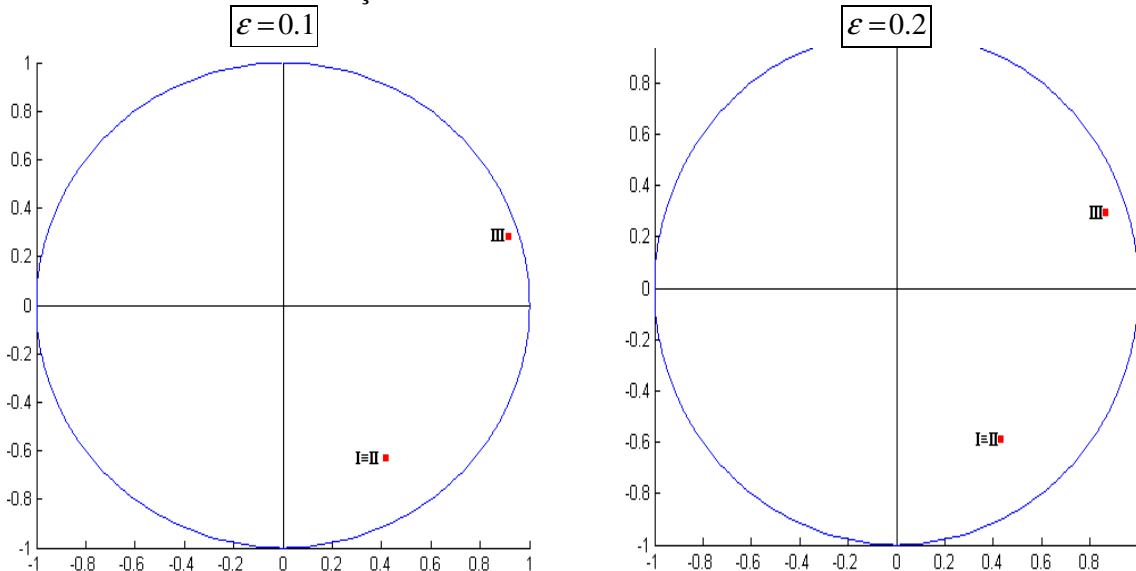


Fig. 6.18 -Círculo de Correlações (12)

Como era de esperar o bloco com mais peso para a formação do primeiro eixo é o bloco III. Os restantes blocos, com o mesmo número de variáveis, têm igual contributo.

Note-se uma situação já verificada nas matrizes de dois blocos. O bloco III tem sempre correlação positiva com o primeiro eixo. No entanto, o bloco I apresentava correlação negativa na Fig. 6.16 e positiva nas figuras 6.17 e 6.18. Ou seja, quando o bloco de maiores correlações é o bloco superior esquerdo a maior correlação é negativa e quando o bloco de correlações mais fortes é o bloco inferior direito a correlação é positiva.

Muitos outros testes deveriam ser efectuados, em particular variando os valores de ε para cada bloco.

Capítulo VI

Conclusões

Se as dimensões de cada paralelepípedo forem ordenadas as distribuições vão ser, em cada plano, triangulares e, por construção, uniformes, sendo espacialmente o conjunto uma pirâmide triangular (tetraedro irregular) de vértices $A(9,9,9)$, $B(0,0,0)$, $C(0,0,9)$ e $D(0,9,9)$:

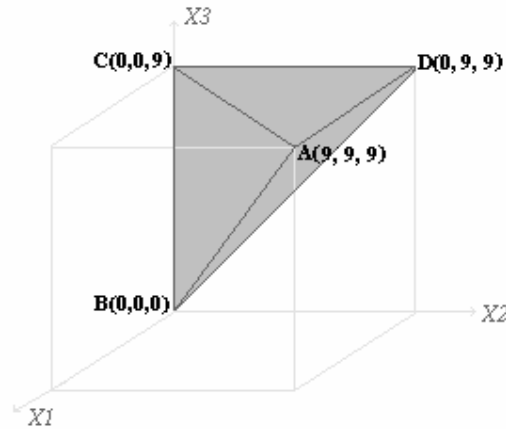


Fig. 7.1 – Geometria dos Paralelepípedos

Se bem se reparar os vértices A e B representam o maior e menor cubo, respectivamente. O vértice C a maior e mais fina ‘tábua’ e o vértice D o ‘palito’ mais fino e mais comprido.

A pirâmide tem seis arestas de comprimentos:

$$\overline{AD} = \overline{BC} = \overline{CD} = 9$$

$$\overline{AC} = \overline{BD} = \sqrt{162} = 9\sqrt{2} \approx 13$$

$$\overline{AB} = \sqrt{243} = 9\sqrt{3} \approx 16$$

A projecção desta pirâmide em qualquer plano vai ser um polígono quadrangular, eventualmente degenerado, se a projecção de algum dos vértices cair no interior ou na aresta do triângulo definido pelos outros três.

Esta geometria justifica a estrutura triangular obtida nas diversas representações dos paralelepípedos (e nas suas variantes) no plano formado pelos dois eixos principais de inércia. Em particular, no exemplo a quatro dimensões, a estrutura pentagonal explica-se pela estrutura espacial dos paralelepípedos ser uma

hiper pirâmide triangular com 5 vértices, e portanto, a sua projecção no plano ser um pentágono se nenhum vértice se projectar no interior do triângulo.

Por outro lado, e mais importante, se se pensar na geometria do espaço amostral (pirâmide triangular) verifica-se que forma e tamanho estão expressos desde o princípio. Assim, e apesar de repetido por diversos autores como Cooley & Lohnes, Davis, FC-FS, o problema de forma e tamanho está claro desde os dados: medidas das arestas.

Juntar variáveis no caso dos paralelepípedos parece só vir aumentar a entropia, pois as estruturas fundamentais (tábuas, palitos, cubos) estão presentes desde o início, reconfortando todas as técnicas ligadas à diminuição do número de variáveis (escolha de subconjuntos de variáveis) empregues por autores como Cadima, Duarte Silva, Jolliffe, Rousson e Vines.

De facto, as componentes principais, neste caso, parecem não trazer qualquer nova informação. Correspondem às maiores variações e, portanto, seguem arestas.

A primeira componente está, naturalmente, ligada à maior aresta, $\overline{AB}$, e opõe pequenos e grandes cubos (Fig. 7.2.). A segunda componente opõe todas as dimensões iguais a duas muito diferentes, e está ligada às arestas intermédias (Fig. 7.2.):

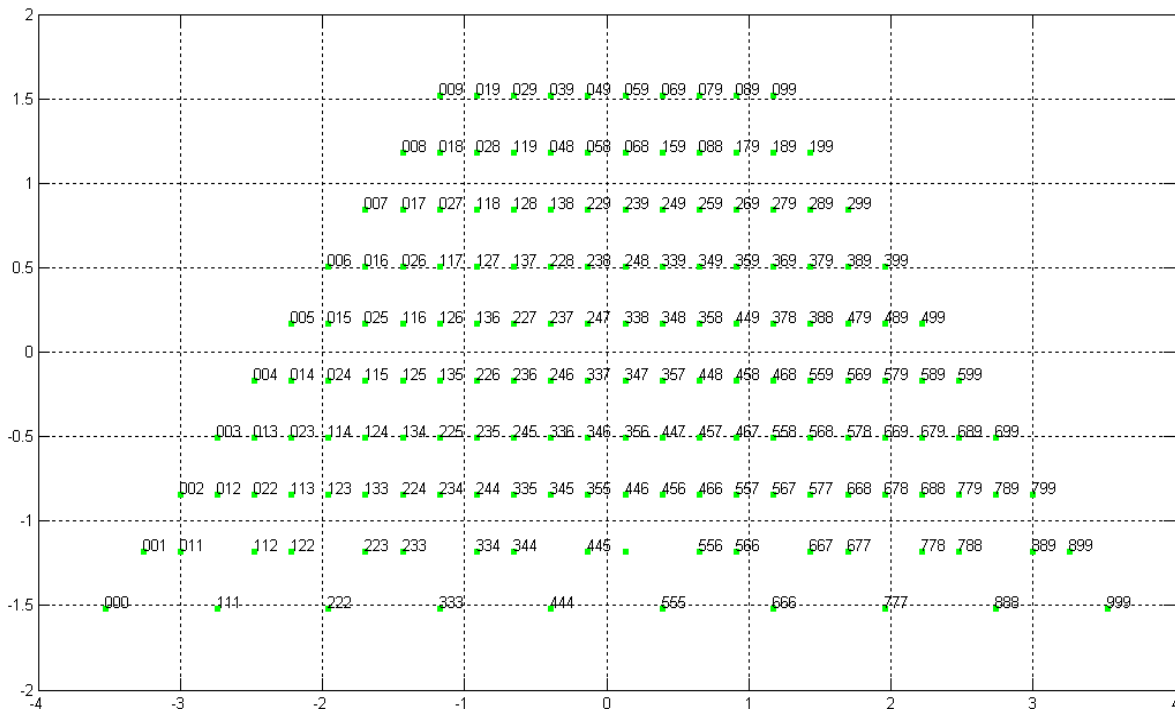


Fig. 7.2. - Indivíduos representados no plano principal (1,2)

A terceira componente opõe tábuas e palitos e está ligada à aresta $\overline{CD}$ (Fig. 7.3.):

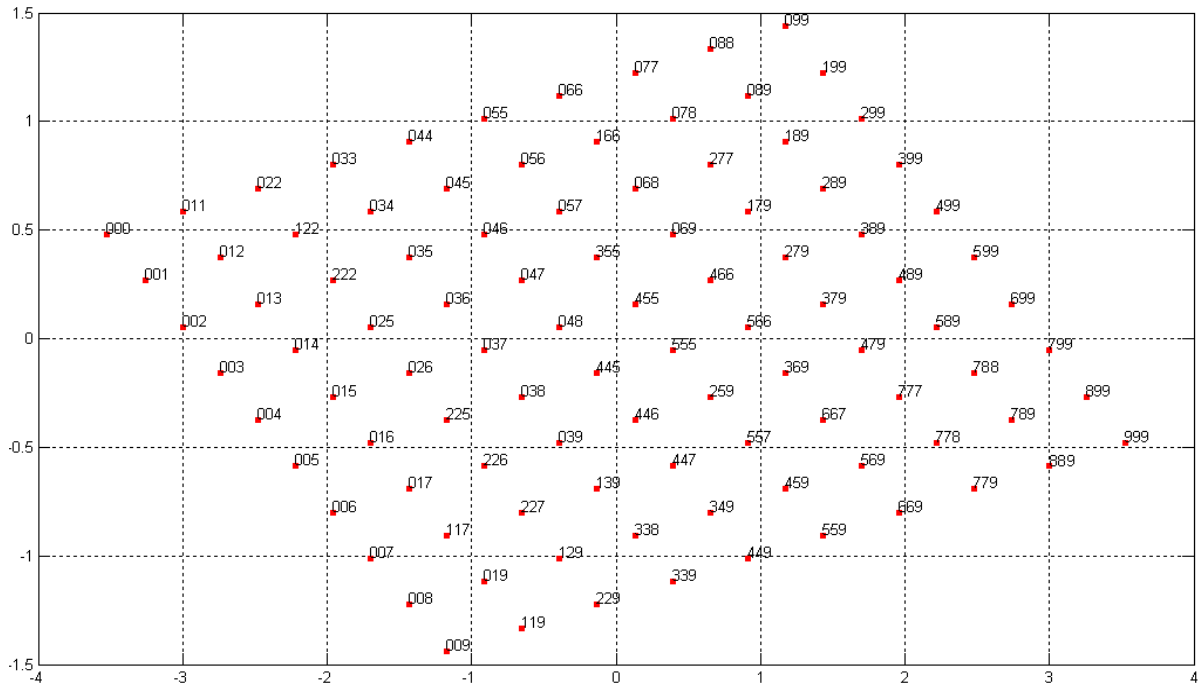


Fig. 7.3. - Indivíduos representados no plano principal (1,3)

As variáveis construídas vão privilegiar vértices. Os polinómios de maior potência privilegiam os grandes volumes e as grandes áreas: vértice A.

As variáveis $Máximo - Mínimo$ e $\frac{Máximo}{Mínimo}$ privilegiam diferenças para a cubicidade.

Pese embora o que foi referido anteriormente, o exemplo dos paralelepípedos continua a ser um bom exemplo para testar teorias, visto ser completamente geométrico.

Pode também verificar-se que as variáveis se organizam em blocos na matriz de correlações, sendo que essa organização se processa, não pelo tipo de informação subjacente a cada uma das variáveis, mas pelo tipo matemático em que se incluem. A dimensão destes blocos está, directamente, relacionada com as componentes principais na medida em que as variáveis que mais contribuem para a formação da primeira componente são as variáveis pertencentes ao bloco de maior dimensão. Analogamente, para a formação da segunda componente são as variáveis pertencentes ao segundo maior bloco que apresentam maior contributo.

No caso dos paralelepípedos as variáveis consideradas dividem-se em: analíticas (maior bloco), funções dos extremos (bloco intermédio) e máximo. Assim, a primeira componente é analítica porque as

variáveis analíticas são em maior número e a segunda componente está mais correlacionado com os extremos pois constituem o segundo bloco de maior dimensão.

Deste modo, a ACP revelou-se facilmente manipulável, pois é a escolha das variáveis que fixa as componentes. Basta juntar ou subtrair variáveis, convenientemente, para que determinado tipo de informação tenha mais peso.

Para os diversos testes realizados em termos de ordenação das arestas dos paralelepípedos observa-se que ordenar ou desordenar os dados não interfere nas primeiras componentes pois a dominância do primeiro bloco mantém-se, embora a ordenação possa implicar a perda de alguma informação caso exista uma estrutura subjacente: temporal no caso das Águas de Gaia, disciplinas no caso dos alunos. Contudo, se o objectivo passar por uma mera destrição em termos de ‘tamanho’ a questão da ordenação é totalmente irrelevante.

Se as dimensões não estiverem ordenadas as componentes vão ser meras consequências das flutuações da amostragem: dados completamente simétricos em relação aos eixos. Para o conjunto de todos os paralelepípedos a variância explicada é nula. As direcções principais são as direcções principais de uma nuvem cúbica.

Pode também admitir-se que o caso discreto é uma boa descrição do caso contínuo, pois considerando o caso contínuo as conclusões são idênticas ao discreto quer em termos de interpretação da ACP, quer em termos de estabilidade.

BIBLIOGRAFIA

- [13] Arvazian, S. (1978),
'*Étude Statistique des Dépendences*', Edu. Mir-Moscovo, 2ªEdição.
- [4] Cadima, J.F. & Jolliffe, I.T (1995),
'Loadings and Correlations in the Interpretation of Principal Components', Journal of Applied Statistics **22** (2), 203-214.
- [6] Cadima, J.F. & Jolliffe, I.T (2001),
'Variable Selection and the Interpretation of Principal Subspaces', Journal of Agricultural, Biological and Environmental Statistics **6** (1), 62-79.
- [8] Cadima, J.F, Cerdeira, J.O & Minhoto, P. (2004),
'Computational aspects of algorithms for variable selection in the context of principal components', Computational Statistics & Data Analysis **47**, 225-236.
- [1] Calheiros, F. & Sousa, F. (1991),
'Robustez em Análises em Componentes Principais de Variáveis Não Gaussianas', Separata das 1ªs Jornadas de Estatística e Aplicações, 114-126, Universidade do Minho.
- [3] Cooley, W.W & Lohnes, P.R (1971)
'*Multivariate Data Analysis*', John Wiley & Sons.
- [2] Davis, J. (1977)
'*Statistics and Data Analysis in Geology*', Edu. Wiley.
- [7] Duarte Silva, A.P. (2002),
'Discarding Variables in a Principal Component Analysis: Algorithms for All Subsets Comparisons', Comput. Statist. **17**, 251-271.
- [12] Figueiredo, A.
Tese da Universidade de Ciências da Universidade de Lisboa
- [11] Gomes, P.(1987),
'*Distribution de Bingham sur la n-sphere: une nouvelle approche de l'analyse factorielle*', Thèse d'État, Montpellier.
- [14] Greenacre, M.J. (2004)
'*Theory and Applications of Correspondence Analysis*', Academic Press.
- [4] McCabe, G.P (1975)
'*Principal Variables*', Technometrics **26**, 137-144.
- [11] Mardia, K. V. & Jupp, P. E (1999).
'*Directional Statistics*', John Wiley & Sons, Londres.
- [9] Vines, S. K. (2000),
'*Simple principal Components*', Applied Statistics **49**, 441-451.
- [10] Rousson, V. & Gasser, T. (2004),
'*Simple Component Analysis*', Applied Statistics **53**(4), 539-555.
- [15] Saporta, G. (1990)
'*Probabilités Analyse des Données et Statistique*', Edu. Technip, Paris.

ANEXOS

Anexo I

Perfis de Consumidores

O estudo sobre perfis de consumidores, cujo resumo se apresenta, inseriu-se no âmbito da cadeira de seminário do Mestrado em Estatística Aplicada e Modelação.

Como ponto de partida utilizou-se um ficheiro de dados bimensais do consumo de água da cidade de Vila Nova de Gaia, referentes ao período compreendido entre Janeiro de 1999 e Fevereiro de 2003, num total de 10090251 registos.

Para além da informação sobre o consumo, o ficheiro continha a identificação numérica do cliente, informação sobre a freguesia e morada, reservatório correspondente, bem como a facturação e data de facturação.

Contudo, na cadeira de Seminário foram apenas analisados os dados relativos a um dos reservatórios considerados no ficheiro de dados inicial. Em particular, são analisados os dados referentes ao Reservatório 19, o que equivale a um total de 57706 registos.

Foi feita uma primeira análise ao número total de registos por consumidor que permitiu verificar não serem em igual número para todos os consumidores. Por outro lado, analisou-se com algum cuidado o comportamento do consumidor em termos de contagens: número de contagens em cada ano, número de contagens em cada mês considerado e número de contagens por mês. Verificou-se que não obstante a grande maioria dos consumidores ter apenas contagens nos bimeses existiam consumidores com registos para todos os meses do ano e outros com registos esporádicos entre os bimeses. Observou-se também a existência de consumidores com mais do que uma contagem por mês (esta situação poder-se-á dever a contagens de correcção ou a erros do próprio ficheiro).

Quanto à organização dos dados, os consumos foram ordenados cronologicamente para cada consumidor.

Findo este primeiro contacto com os dados, passou-se a uma análise mais exaustiva recorrendo ao software *MatLab*.

O primeiro passo consistiu na escolha das variáveis. Tratou-se de uma escolha experimental, não formalizada.

Definiram-se inicialmente seis variáveis para caracterizar cada consumidor. No entanto, no decorrer do trabalho novas variáveis foram sendo acrescentadas. Estas variáveis surgiram quer por necessidade de novas informações capazes de caracterizar os consumidores quer por sugestão da metodologia aplicada.

Foram, finalmente, consideradas 19 variáveis:

- Variáveis 1 a 6 – *Índices Sazonais* (Jan/Mar/Mai/Jul/Set/Nov)
- Variável 7 – *Tendência* (T)
- Variável 8 – *Número de contagens* (NC)
- Variável 9 – *Número de contagens corrigidas* (NCC)
- Variável 10 – *Perfil do consumidor* (PC)
- Variável 11 – *Perfil mediano do consumidor* (PMC)
- Variável 12 – *Perfil actual do consumidor* (PA)
- Variável 13 – *Perfil mediano actual do consumidor* (PMA)
- Variável 14 – *Coefficiente de variação* (CV)
- Variável 15 – *Máximo* (Max)
- Variável 16 – *Mínimo* (Min)
- Variável 17 – *Mínimo – Máximo* (MaxMin)
- Variável 18 – *Declive da recta dos mínimos quadrados* (D)
- Variável 19 – *Ordenada na origem da recta dos mínimos quadrados* (OO)

No estudo preliminar tinha-se verificado que o número de contagens diferia em alguns consumidores, pelo que se tornou necessário criar uma nova variável representativa do número de contagens por consumidor: variável *Número de Contagens*.

Por outro lado, alguns consumidores apresentavam mais do que uma contagem em determinados meses. Note-se que em grande parte dos casos as contagens excedentes representam um consumo zero. A solução encontrada consistiu em considerar o acumulado do consumo nas diferentes contagens do mês em questão.

Um dos problemas encontrados no tratamento dos dados do ficheiro foi a presença de consumos negativos, cuja resolução foi efectuada através de correcção democrática para trás. Ou seja, o primeiro consumo negativo é corrigido calculando o consumo acumulado até à contagem negativa (inclusive) e dividindo-o pelo respectivo número de contagens consideradas. Essa quantidade é distribuída equitativamente pelas células a serem corrigidas (sendo susceptível de ser alterada, por forma a que os acumulados se verifiquem). A correcção entre consumos negativos é feita do mesmo modo, mas considerando a diferença de acumulados entre valores negativos. Nos casos em que a primeira contagem disponível é negativa considerou-se o valor correspondente ao mesmo mês do ano seguinte.

Assim, dada a necessidade da correcção de alguns dos valores disponíveis julgou-se relevante criar uma variável representativa do número de correcções efectuadas: *Número de contagens corrigidas*.

A variável **Tendência** pretende avaliar a evolução dos consumidores. Nesse sentido é calculada a tendência utilizando o método das Médias Móveis.

Após o cálculo da variável tendência passou-se à determinação das variáveis respeitantes à sazonalidade. Os índices sazonais são calculados efectuando a média dos consumos sem tendência em cada mês.

Aquando do cálculo da tendência foram apenas calculadas a primeira e última médias móveis pois a determinação de todas as médias móveis tornar-se-ia computacionalmente pesada.

Assim, para poder calcular os índices sazonais foi necessário estimar a tendência. Para tal foi usada a recta dos mínimos quadrados:

$$y = mx + b$$

onde o declive $m = \frac{\text{cov}(X, Y)}{\sigma^2_x}$ e a ordenada na origem $b = \bar{Y} - m\bar{X}$. Note-se que o vector Y representa os consumos de cada consumidor (ordenados cronologicamente) e X representa o vector [1, 2, ..., dim (Y)].

O declive e ordenada da origem correspondentes a cada consumidor foram também considerados como variáveis: **Declive da recta dos mínimos quadrados** e **Ordenada na origem da recta dos mínimos quadrados**. O declive caracteriza a tendência dos mínimos quadrados sendo que o seu valor juntamente com o valor da ordenada na origem permitem retirar algumas conclusões quanto ao comportamento do consumidor (por exemplo, um declive pequeno poderá indicar um consumidor regular cujo consumo médio estará próximo do valor da ordenada na origem).

Denominou-se **Perfil do consumidor** a variável que representa o consumo médio de cada consumidor e **Perfil mediano do consumidor** a variável representativa da mediana dos consumos.

O **Perfil actual do consumidor** é calculado através média dos seis últimos consumos, enquanto o **Perfil mediano actual do consumidor** é dado pela mediana dos últimos seis consumos de cada consumidor.

O **Coefficiente de variação** é calculado da seguinte forma: $\frac{PC - PM}{PC}$ para cada consumidor.

Aplicou-se análise em componentes principais à matriz de dados obtida.

A primeira ACP foi realizada considerando como matriz de pesos a matriz identidade.

Conclui-se que alguns indivíduos eram muito preponderantes sobre os restantes. Além disso o número de contagens não é o mesmo para todos os consumidores. Considerou-se, assim, ser vantajoso pesar os indivíduos.

A matriz de pesos P considerada é tal que:

$$p_{ii} = \frac{\text{Número de contagens do consumidor } i}{\text{Número Total de Conatgens}}, \quad i = 1..n^\circ \text{ de consumidores considerados}$$

Verificou-se que em relação à ACP realizada com os indivíduos devidamente pesados não foram significativas.

Para estas variáveis obteve-se a seguinte matriz de correlações:

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19
X1	1	0,3	0,19	-0,08	-0,65	0,21	-0,92	-0,12	-0,05	-0,63	-0,69	-0,74	-0,72	-0,01	-0,66	0,04	0,68	-0,85	-0,3
X2	0,3	1	-0,51	0,04	0,13	0,72	-0,37	-0,05	-0,02	-0,02	-0,01	-0,16	-0,14	0,02	-0,05	0,11	0,06	-0,39	0,24
X3	0,19	-0,51	1	0,11	-0,58	-0,36	-0,16	0,04	0,01	-0,62	-0,6	-0,5	-0,54	0	-0,58	-0,03	0,59	-0,2	-0,72
X4	-0,08	0,04	0,11	1	0,37	0,36	0,11	0,24	0,01	0,33	0,32	0,32	0,31	-0,01	0,28	0,55	-0,25	0,15	0,38
X5	-0,65	0,13	-0,58	0,37	1	0,33	0,71	0,13	0,02	0,82	0,87	0,82	0,82	0	0,8	0,24	-0,81	0,73	0,67
X6	0,21	0,72	-0,36	0,36	0,33	1	-0,23	0	-0,02	0,03	0,05	-0,07	-0,06	0,01	-0,04	0,17	0,05	-0,17	0,2
X7	-0,92	-0,37	-0,16	0,11	0,71	-0,23	1	0,1	0,04	0,69	0,76	0,8	0,78	0,01	0,69	0,1	-0,7	0,96	0,31
X8	-0,12	-0,05	0,04	0,24	0,13	0	0,1	1	0,06	0,15	0,14	0,16	0,15	0,01	0,18	0,09	-0,17	0,1	0,14
X9	-0,05	-0,02	0,01	0,01	0,02	-0,02	0,04	0,06	1	0,03	0,03	0,03	0,03	0,13	0,03	-0,02	-0,03	0,04	0,01
X10	-0,63	-0,02	-0,62	0,33	0,82	0,03	0,69	0,15	0,03	1	0,99	0,98	0,99	-0,01	0,98	0,46	-0,97	0,73	0,9
X11	-0,69	-0,01	-0,6	0,32	0,87	0,05	0,76	0,14	0,03	0,99	1	0,98	0,99	-0,01	0,96	0,43	-0,95	0,79	0,85
X12	-0,74	-0,16	-0,5	0,32	0,82	-0,07	0,8	0,16	0,03	0,98	0,98	1	1	-0,01	0,96	0,41	-0,96	0,82	0,81
X13	-0,72	-0,14	-0,54	0,31	0,82	-0,06	0,78	0,15	0,03	0,99	0,99	1	1	-0,01	0,96	0,4	-0,96	0,8	0,83
X14	-0,01	0,02	0	-0,01	0	0,01	0,01	0,01	0,13	-0,01	-0,01	-0,01	-0,01	1	0	-0,08	0	0	-0,02
X15	-0,66	-0,05	-0,58	0,28	0,8	-0,04	0,69	0,18	0,03	0,98	0,96	0,96	0,96	0	1	0,39	-1	0,71	0,87
X16	0,04	0,11	-0,03	0,55	0,24	0,17	0,1	0,09	-0,02	0,46	0,43	0,41	0,4	-0,08	0,39	1	-0,35	0,16	0,55
X17	0,68	0,06	0,59	-0,25	-0,81	0,05	-0,7	-0,17	-0,03	-0,97	-0,95	-0,96	-0,96	0	-1	-0,35	1	-0,72	-0,86
X18	-0,85	-0,39	-0,2	0,15	0,73	-0,17	0,96	0,1	0,04	0,73	0,79	0,82	0,8	0	0,71	0,16	-0,72	1	0,37
X19	-0,3	0,24	-0,72	0,38	0,67	0,2	0,31	0,14	0,01	0,9	0,85	0,81	0,83	-0,02	0,87	0,55	-0,86	0,37	1

Tabela 1

Reorganiza-se a matriz em blocos de correlações, de forma que em cada bloco todas as correlações sejam superiores, em módulo, a 50%. Obviamente que esta opção pode levar a que os blocos não estejam totalmente destrinchados, sendo que nas antidiagonais podem existir também correlações significativas, mas sempre em menor número:

	X1	X7	X18	X12	X13	X11	X15	X17	X5	X10	X2	X6	X3	X19	X16	X4	X8	X9	X14
X1	1	-0,92	-0,85	-0,74	-0,72	-0,69	-0,66	0,68	-0,65	-0,63	-0,37	-0,23	-0,16	0,31	0,1	0,11	0,1	0,04	0,01
X7	-0,92	1	0,96	0,8	0,78	0,76	0,69	-0,7	0,71	0,69	-0,02	0,03	-0,62	0,9	0,46	0,33	0,15	0,03	-0,01
X18	-0,85	0,96	1	0,82	0,8	0,79	0,71	-0,72	0,73	0,73	-0,01	0,05	-0,6	0,85	0,43	0,32	0,14	0,03	-0,01
X12	-0,74	0,8	0,82	1	1	0,98	0,96	-0,96	0,82	0,98	-0,16	-0,07	-0,5	0,81	0,41	0,32	0,16	0,03	-0,01
X13	-0,72	0,78	0,8	1	1	0,99	0,96	-0,96	0,82	0,99	-0,14	-0,06	-0,54	0,83	0,4	0,31	0,15	0,03	-0,01
X11	-0,69	0,76	0,79	0,98	0,99	1	0,96	-0,95	0,87	0,99	-0,05	-0,04	-0,58	0,87	0,39	0,28	0,18	0,03	0
X15	-0,66	0,69	0,71	0,96	0,96	0,96	1	-1	0,8	0,98	0,06	0,05	0,59	-0,86	-0,35	-0,25	-0,17	-0,03	0
X17	0,68	-0,7	-0,72	-0,96	-0,96	-0,95	-1	1	-0,81	-0,97	-0,39	-0,17	-0,2	0,37	0,16	0,15	0,1	0,04	0
X5	-0,65	0,71	0,73	0,82	0,82	0,87	0,8	-0,81	1	0,82	0,3	0,21	0,19	-0,3	0,04	-0,08	-0,12	-0,05	-0,01
X10	-0,63	0,69	0,73	0,98	0,99	0,99	0,98	-0,97	0,82	1	0,13	0,33	-0,58	0,67	0,24	0,37	0,13	0,02	0
X2	0,3	-0,37	-0,39	-0,16	-0,14	-0,01	-0,05	0,06	0,13	-0,02	1	0,72	-0,51	0,24	0,11	0,04	-0,05	-0,02	0,02
X6	0,21	-0,23	-0,17	-0,07	-0,06	0,05	-0,04	0,05	0,33	0,03	0,72	1	-0,36	0,2	0,17	0,36	0	-0,02	0,01
X3	0,19	-0,16	-0,2	-0,5	-0,54	-0,6	-0,58	0,59	-0,58	-0,62	-0,51	-0,36	1	-0,72	-0,03	0,11	0,04	0,01	0
X19	-0,3	0,31	0,37	0,81	0,83	0,85	0,87	-0,86	0,67	0,9	0,24	0,2	-0,72	1	0,55	0,38	0,14	0,01	-0,02
X16	0,04	0,1	0,16	0,41	0,4	0,43	0,39	-0,35	0,24	0,46	0,11	0,17	-0,03	0,55	1	0,55	0,09	-0,02	-0,08
X4	-0,08	0,11	0,15	0,32	0,31	0,32	0,28	-0,25	0,37	0,33	0,04	0,36	0,11	0,38	0,55	1	0,24	0,01	-0,01
X8	-0,12	0,1	0,1	0,16	0,15	0,14	0,18	-0,17	0,13	0,15	-0,05	0	0,04	0,14	0,09	0,24	1	0,06	0,01
X9	-0,05	0,04	0,04	0,03	0,03	0,03	0,03	-0,03	0,02	0,03	-0,02	-0,02	0,01	0,01	-0,02	0,01	0,06	1	0,13
X14	-0,01	0,01	0	-0,01	-0,01	-0,01	0	0	0	-0,01	0,02	0,01	0	-0,02	-0,08	-0,01	0,01	0,13	1

Tabela 2

Observa-se então que ao maior bloco pertencem todas as variáveis de Perfil, Tendência, Declive da recta dos mínimos quadrados, Índices Sazonais referentes a Janeiro e Setembro, Máximo e a diferença Mínimo – Máximo. Note-se que das variáveis no bloco antidiagonal a variável *Ordenada na origem da recta dos mínimos quadrados* está muito correlacionada com todas as variáveis de perfil e extremos.

Pode formar-se um segundo bloco constituído pelos índices sazonais referentes a Março e Novembro. As correlações com as restantes variáveis são, à excepção da correlação com o índice sazonal de Maio, inferiores a 50%.

A variável 19 embora tenha, como foi referido, fortes correlações com algumas das variáveis do primeiro bloco, pode incluir-se juntamente com o índice sazonal referente a Março num outro bloco.

Mínimo e Julho inserem-se num 4º bloco, sendo todas as suas correlações com as restantes variáveis pouco significativas.

Número de contagens, Número de contagens corrigidas e Coeficiente de variação apresentam-se isoladas.

O quadro seguinte representa a quantidade de informação explicada pelos eixos principais:

Valores Próprios	Inércia Explicada (%)	Inércia Explicada (acumulada) (%)
9,6891	50,995	50,9951
2,7885	14,676	65,6715
1,5907	8,372	74,0435
1,1726	6,171	80,2149
1,0359	5,452	85,6670
0,8582	4,517	90,1838
0,9087	4,782	94,9661
0,3971	2,090	97,0562
0,2442	1,285	98,3416
0,1108	0,583	98,9248
0,1041	0,548	99,4729
0,0582	0,306	99,7792
0,0257	0,135	99,9145
0,0126	0,066	99,9807
0,0029	0,015	99,9958
0,0005	0,003	99,9986
0,0002	0,001	99,9996
0,0001	0,000	100
0,0000	0,000	100

Tabela 3

Serão neste caso retidas 4 componentes. Esta escolha baseia-se no facto de as quatro primeiras componentes principais explicarem cerca de 80% da inércia total.

A primeira componente explica a maior parte da inércia: 51%. A segunda componente explica 15% enquanto que a terceira e quarta componentes explica 8% e 6% da informação total, respectivamente.

Pôde verificar-se que para a construção do 1º componente principal não existe uma variável cujo contributo se evidencie em relação aos restantes. As variáveis de perfil contribuem de forma muito idêntica seguidas de perto pelas variáveis Máximo e Mínimo-Máximo. É importante realçar que todas estas pertencem ao maior bloco de correlações.

Para a 2ª componente as variáveis que mais contribuem são Março e Novembro (variáveis de Sazonalidade), pertencentes ao segundo bloco da diagonal.

Quanto à 3ª componente a variável com maior contributo é Julho, embora as variáveis Mínimo e Maio também contribuam significativamente. Estas variáveis não pertencem a um só bloco, mas aos 3º e 4º blocos da diagonal. Provavelmente, porque dado o menor peso desta componente a informação se encontra mais diluída pelas variáveis e não se relaciona tão fortemente apenas com um bloco.

Para o 4º eixo contribuem, em pé de igualdade, as variáveis Coeficiente de Variação e Número de Contagens Corrigidas, variáveis isoladas em termos de blocos.

Segue-se a representação dos consumidores na primeira e segunda componentes principais:

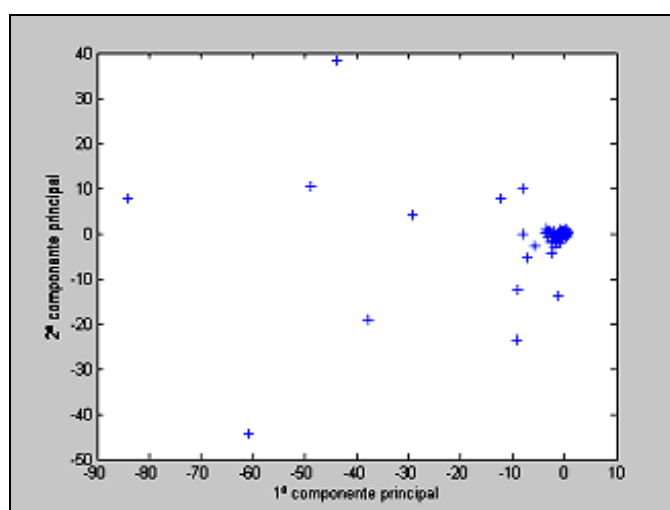


Fig. 1 - Consumidores representados na 1ª e 2ª componentes principais

Existe um grupo de indivíduos que contribuem de forma efectiva para a formação de todos os eixos. São os consumidores que apresentam elevados durante todo o ano. Tratam-se muito provavelmente, e tendo em conta o mapa do reservatório considerados, de caves do vinho do Porto (consumidores institucionais).

Existe um pequeno grupo intermédio e perto da origem uma nuvem de pontos dos quais não é possível destacar qualquer consumidor. Este último grupo corresponde a consumidores tradicionais com registos de consumo muito inferiores aos apresentados pelos consumidores mencionados.

Dado o peso assinalável dos consumidores referenciados como institucionais torna-se interessante verificar o que ocorre efectuando a ACP sem que estes sejam considerados.

Aliás, uma hipótese considerada para formar grupos de consumidores com determinadas características foi trabalhar em ‘coroa’, ou seja, ir retirando os consumidores que se destacam e ver o que acontece.

Verificou-se que retirando os consumidores institucionais a matriz de correlações sofre bastantes alterações.

Observa-se que a maioria das correlações se tornam menos significativas. Há contudo variáveis que anteriormente se apresentavam muito pouco correlacionadas com as restantes e que após a eliminação dos consumidores ganham correlações mais significativas.

A variável Setembro, antes da retirada dos indivíduos, apresentava elevada correlação com as variáveis de perfil. Essas correlações são agora pouco significativas. Esta situação deve-se ao elevado consumo apresentado pelos consumidores eliminados, em particular no mês de Setembro. Este elevado consumo no mês de Setembro confirma as suspeitas que se poderão tratar de caves do vinho do Porto dado que Setembro se trata de mês de vindimas.

Pode inclusivamente verificar-se a alteração do sinal de algumas correlações, em particular nas variáveis de sazonalidade. Para a nova matriz de correlações os meses de Verão correlacionam-se positivamente e negativamente com os meses de Inverno. Na matriz inicial era claro o sinal da correlação entre meses de Verão e entre meses de Inverno, não existindo acordo em relação ao sinal da correlação entre meses de Verão e Inverno. A justificação poderá dever-se ao facto de dado o elevado registo de consumo para estes consumidores ao longo de todo o ano as diferenças entre meses de Verão e Inverno serem atenuadas.

Assim, realizando a ACP sobre os dados foi possível distinguir um grupo de consumidores com consumos claramente superiores aos registados pelos restantes consumidores. São, provavelmente, caves de vinho do Porto.

Retirando estes consumidores surge um novo grupo de consumidores com consumo intermédio.

Por outro lado, foi possível verificar que conjugando a informação das diversas variáveis se podem retirar várias informações quanto ao comportamento dos consumidores ou, pelo menos, formular algumas hipóteses.

Apesar do estudo realizado retirando variáveis não ter sido muito exaustivo o efectuado não acrescentou muita informação.

Anexo II

Documentação do Programa

Este programa tem como objectivo gerar todos os diferentes paralelepípedos com coordenadas inteiras de 1 a 10.

Foi testado aplicando um exemplo presente em Saporta [15] (Página 180).

Foram construídos os seguintes M-Files:

Paralelepipedos
[X,Media,V]=standarizacao(X)
ValoresPropriosMatrizCovariancias
Componentes
Circulocorrelacoes

M-FILE: *Parelelepípedos.m*

São gerados paralelepípedos cujas dimensões tomam todos os valores inteiros no intervalo [1,10] e colocados na matriz X; assim cada coordenada vai ter as 3 coordenadas do paralelepípedo.

Esta geração foi dividida em três etapas. Numa primeira etapa, são gerados todos os cubos possíveis com as dimensões consideradas. Numa segunda etapa, partindo dos cubos obtidos, foram gerados prismas quadrangulares. Finalmente, obtiveram-se os paralelepípedos.

Segue-se o programa dividido nas três diferentes etapas:

Cubos	<pre> for i=1:10 Y(i,:)=i,i,i; end </pre>
Prismas Quadrangulares	<pre> X=[]; for i=1:10 for j=1:10 X1(j,1:2)=Y(i,1:2); X1(j,3)=j; end X=[X;X1]; end </pre>
Paralelepípedos	<pre> X=[X; nchoosek(1:10,3)]; </pre>

As dimensões dos paralelepípedos obtidos são ordenadas por ordem crescente:

**Ordenação
das Dimensões**

```
X=sort(X,2);
```

Após a geração dos paralelepípedos procedeu-se ao cálculo das variáveis pretendidas:

Cálculo das Variáveis

```
a=length(X)

X12=X(:,1).*X(:,2);
X13=X(:,1).*X(:,3);
X23=X(:,2).*X(:,3);

X123=[X12,X13,X23];

X(:,4)=X(:,1).*X(:,2).*X(:,3);
X(:,5)=2*sum(X123,2);

for i=1:a
X(i,6)=max(X(i,1:3));
end

for i=1:a
X(i,7)=min(X(i,1:3));
end
for i=1:a
X(i,8)=mean(X(i,1:3));
end

for i=1:a
X(i,9)=mean(X123(i,1:3));
end

X(:,10)=X(:,6)./X(:,7);
X(:,11)=X(:,6)-X(:,7);

for i=1:a
X(i,12)=min(X(i,1:3))/max(X(i,1:3));
end
```

A matriz X final é obtida permutando as linhas entre si, para que os paralelepípedos se encontrem ‘desordenados’.

```
Trocar=randperm(220);
X=X(Trocar,:);
```

M-FILE: Estandarizacao.m

A função seguinte permite centrar (média zero) e reduzir (variância um) a matriz de dados:

Dados Centrados e Reduzidos	<pre>function[X,Media,V]=standarizacao(X) n=size(X,1); Media=mean(X); Media=repmat(Media,n,1); X=X-Media; V=cov(X); d=diag(1./sqrt(diag(V))); X=X*d; Media=Mean(X); V=Cov(X); end</pre>
------------------------------------	---

M-FILE: ValoresPropriosMatrizCovariancias.m

No bloco abaixo segue o cálculo dos valores próprios e vectores próprios associados (normalizados) da matriz de correlações. Em particular, os vectores próprios formam as colunas da matriz U e os valores próprios a diagonal principal da matriz (diagonal) A.

Cálculo dos Valores e Vectores Próprios	<pre>[U,A]=eig(V); for i=1:size(U,2) a=normest(U(:,i)); U(:,i)=U(:,i)/a; end A=diag(A); a=size(X,2);</pre>
--	---

Ordenação dos Valores Próprios

```
if A(1)~=max(A)
    A=sort(A,'descend')
    U=U(:,a:-1:1)
end
```

M-FILE: Componentes.m

As componentes principais constituem a matriz Y, obtida como se indica abaixo.

Cálculo das Componentes Principais Representação dos indivíduos nos 1º, 2º e 3º eixos principais

```
Y=X*U

figure
subplot(1,3,1);
plot(Y(:,1),Y(:,2),'rs','MarkerSize',4,'MarkerEdgeColor','g','MarkerFaceColor','g')
title 'Primeira e Segunda Componentes Principais'
grid on

subplot(1,3,2);
plot(Y(:,1),Y(:,3),'rs','MarkerSize',4,'MarkerEdgeColor','r','MarkerFaceColor','r')
title 'Primeira e Terceira Componentes Principais'
grid on

subplot(1,3,3)
plot(Y(:,2),Y(:,3),'rs','MarkerSize',4,'MarkerEdgeColor','b','MarkerFaceColor','b')
title 'Segunda e Terceira Componentes Principais'
grid on

pause
gname
```

M-FILE: correlacoes.m

Construção do círculo de correlações:

```
a=size(X,2);

for j=1:a
    b=sqrt(A(j));
    r(:,j)=b*U(:,j);
end

figure
subplot(1,3,1);

for k=0:1
```

```

        y=-1:0.001:1;
        recta=tan(k*pi/2)*y;
        hold on
        plot(y,recta,'Color','black')
        axis ([-1 1 -1 1])
    end

    hold on
    y=-1:0.001:1;
    t = 0:pi/100:2*pi;

    plot(sin(t),cos(t),r(:,1),r(:,2),'rs','MarkerSize',4,'MarkerEdgeColor','g','MarkerFaceColor','g')
    axis ([-1 1 -1 1])
    title 'Primeira e Segunda Componentes Principais'

    subplot(1,3,2);
    for k=0:1
        y=-1:0.001:1;
        recta=tan(k*pi/2)*y;
        hold on
        plot(y,recta,'Color','black')
        axis ([-1 1 -1 1])
    end

    hold on
    y=-1:0.001:1;
    t = 0:pi/100:2*pi;

    plot(sin(t),cos(t),r(:,1),r(:,3),'rs','MarkerSize',4,'MarkerEdgeColor','r','MarkerFaceColor','r')
    axis ([-1 1 -1 1])
    title 'Primeira e Terceira Componentes Principais'

    subplot(1,3,3);
    for k=0:1
        y=-1:0.001:1;
        recta=tan(k*pi/2)*y;
        hold on
        plot(y,recta,'Color','black')
        axis ([-1 1 -1 1])
    end

    hold on
    y=-1:0.001:1;
    t = 0:pi/100:2*pi;

    plot(sin(t),cos(t),r(:,2),r(:,3),'rs','MarkerSize',4,'MarkerEdgeColor','b','MarkerFaceColor','b')
    axis ([-1 1 -1 1])
    title 'Segunda e Terceira Componentes Principais'

    gname

```

Foram criados outros M-Files, não apresentados, em particular para o estudo da estabilidade e estudo dos paralelepípedos contínuos e de quatro dimensões.

ANEXO III**Paralelepípedos J. Davis**

	X1	X2	X3
1	3,76	3,66	0,54
2	8,59	4,99	1,34
3	6,22	6,14	4,52
4	7,57	7,28	7,07
5	9,03	7,08	2,59
6	5,51	3,98	1,3
7	3,27	0,62	0,44
8	8,74	7	3,31
9	9,64	9,49	1,03
10	9,73	1,33	1
11	8,59	2,98	1,17
12	7,12	5,49	3,68
13	4,69	3,01	2,17
14	5,51	1,34	1,27
15	1,65	1,61	1,57
16	5,9	5,76	1,55
17	9,84	9,27	1,51
18	8,39	4,92	2,54
19	4,94	4,38	1,03
20	7,23	2,3	1,77
21	9,46	7,31	1,04
22	9,55	5,35	4,25
23	4,94	4,52	4,5
24	8,21	3,08	2,42
25	9,41	6,44	5,11

ANEXO IV

Paralelepípedos Inteiros

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
047	1	5	8	40	106	8	1	4,6667	17,667	8	7	0,125
344	4	5	5	100	130	5	4	4,6667	21,667	1,25	1	0,8
246	3	5	7	105	142	7	3	5	23,667	2,3333	4	0,42857
037	1	4	8	32	88	8	1	4,3333	14,667	8	7	0,125
268	3	7	9	189	222	9	3	6,3333	37	3	6	0,33333
279	3	8	10	240	268	10	3	7	44,667	3,3333	7	0,3
223	3	3	4	36	66	4	3	3,3333	11	1,3333	1	0,75
688	7	9	9	567	414	9	7	8,3333	69	1,2857	2	0,77778
066	1	7	7	49	126	7	1	5	21	7	6	0,14286
367	4	7	8	224	232	8	4	6,3333	38,667	2	4	0,5
011	1	2	2	4	16	2	1	1,6667	2,6667	2	1	0,5
456	5	6	7	210	214	7	5	6	35,667	1,4	2	0,71429
067	1	7	8	56	142	8	1	5,3333	23,667	8	7	0,125
579	6	8	10	480	376	10	6	8	62,667	1,6667	4	0,6
079	1	8	10	80	196	10	1	6,3333	32,667	10	9	0,1
399	4	10	10	400	360	10	4	8	60	2,5	6	0,4
003	1	1	4	4	18	4	1	2	3	4	3	0,25
089	1	9	10	90	218	10	1	6,6667	36,333	10	9	0,1
012	1	2	3	6	22	3	1	2	3,6667	3	2	0,33333
256	3	6	7	126	162	7	3	5,3333	27	2,3333	4	0,42857
144	2	5	5	50	90	5	2	4	15	2,5	3	0,4
789	8	9	10	720	484	10	8	9	80,667	1,25	2	0,8
888	9	9	9	729	486	9	9	9	81	1	0	1
999	10	10	10	1000	600	10	10	10	100	1	0	1
289	3	9	10	270	294	10	3	7,3333	49	3,3333	7	0,3
009	1	1	10	10	42	10	1	4	7	10	9	0,1
559	6	6	10	360	312	10	6	7,3333	52	1,6667	4	0,6
136	2	4	7	56	100	7	2	4,3333	16,667	3,5	5	0,28571
566	6	7	7	294	266	7	6	6,6667	44,333	1,1667	1	0,85714
288	3	9	9	243	270	9	3	7	45	3	6	0,33333
005	1	1	6	6	26	6	1	2,6667	4,3333	6	5	0,16667
267	3	7	8	168	202	8	3	6	33,667	2,6667	5	0,375
001	1	1	2	2	10	2	1	1,3333	1,6667	2	1	0,5
568	6	7	9	378	318	9	6	7,3333	53	1,5	3	0,66667
119	2	2	10	40	88	10	2	4,6667	14,667	5	8	0,2
024	1	3	5	15	46	5	1	3	7,6667	5	4	0,2
588	6	9	9	486	378	9	6	8	63	1,5	3	0,66667
002	1	1	3	3	14	3	1	1,6667	2,3333	3	2	0,33333
277	3	8	8	192	224	8	3	6,3333	37,333	2,6667	5	0,375
056	1	6	7	42	110	7	1	4,6667	18,333	7	6	0,14286
337	4	4	8	128	160	8	4	5,3333	26,667	2	4	0,5
078	1	8	9	72	178	9	1	6	29,667	9	8	0,11111
678	7	8	9	504	382	9	7	8	63,667	1,2857	2	0,77778
257	3	6	8	144	180	8	3	5,6667	30	2,6667	5	0,375
569	6	7	10	420	344	10	6	7,6667	57,333	1,6667	4	0,6
055	1	6	6	36	96	6	1	4,3333	16	6	5	0,16667

369	4	7	10	280	276	10	4	7	46	2,5	6	0,4
058	1	6	9	54	138	9	1	5,3333	23	9	8	0,11111
159	2	6	10	120	184	10	2	6	30,667	5	8	0,2
129	2	3	10	60	112	10	2	5	18,667	5	8	0,2
147	2	5	8	80	132	8	2	5	22	4	6	0,25
229	3	3	10	90	138	10	3	5,3333	23	3,3333	7	0,3
468	5	7	9	315	286	9	5	7	47,667	1,8	4	0,55556
355	4	6	6	144	168	6	4	5,3333	28	1,5	2	0,66667
477	5	8	8	320	288	8	5	7	48	1,6	3	0,625
459	5	6	10	300	280	10	5	7	46,667	2	5	0,5
799	8	10	10	800	520	10	8	9,3333	86,667	1,25	2	0,8
247	3	5	8	120	158	8	3	5,3333	26,333	2,6667	5	0,375
699	7	10	10	700	480	10	7	9	80	1,4286	3	0,7
689	7	9	10	630	446	10	7	8,6667	74,333	1,4286	3	0,7
138	2	4	9	72	124	9	2	5	20,667	4,5	7	0,22222
177	2	8	8	128	192	8	2	6	32	4	6	0,25
238	3	4	9	108	150	9	3	5,3333	25	3	6	0,33333
235	3	4	6	72	108	6	3	4,3333	18	2	3	0,5
018	1	2	9	18	58	9	1	4	9,6667	9	8	0,11111
123	2	3	4	24	52	4	2	3	8,6667	2	2	0,5
669	7	7	10	490	378	10	7	8	63	1,4286	3	0,7
149	2	5	10	100	160	10	2	5,6667	26,667	5	8	0,2
033	1	4	4	16	48	4	1	3	8	4	3	0,25
029	1	3	10	30	86	10	1	4,6667	14,333	10	9	0,1
259	3	6	10	180	216	10	3	6,3333	36	3,3333	7	0,3
668	7	7	9	441	350	9	7	7,6667	58,333	1,2857	2	0,77778
338	4	4	9	144	176	9	4	5,6667	29,333	2,25	5	0,44444
333	4	4	4	64	96	4	4	4	16	1	0	1
019	1	2	10	20	64	10	1	4,3333	10,667	10	9	0,1
166	2	7	7	98	154	7	2	5,3333	25,667	3,5	5	0,28571
225	3	3	6	54	90	6	3	4	15	2	3	0,5
122	2	3	3	18	42	3	2	2,6667	7	1,5	1	0,66667
017	1	2	8	16	52	8	1	3,6667	8,6667	8	7	0,125
115	2	2	6	24	56	6	2	3,3333	9,3333	3	4	0,33333
335	4	4	6	96	128	6	4	4,6667	21,333	1,5	2	0,66667
111	2	2	2	8	24	2	2	2	4	1	0	1
228	3	3	9	81	126	9	3	5	21	3	6	0,33333
446	5	5	7	175	190	7	5	5,6667	31,667	1,4	2	0,71429
556	6	6	7	252	240	7	6	6,3333	40	1,1667	1	0,85714
113	2	2	4	16	40	4	2	2,6667	6,6667	2	2	0,5
188	2	9	9	162	234	9	2	6,6667	39	4,5	7	0,22222
169	2	7	10	140	208	10	2	6,3333	34,667	5	8	0,2
112	2	2	3	12	32	3	2	2,3333	5,3333	1,5	1	0,66667
069	1	7	10	70	174	10	1	6	29	10	9	0,1
599	6	10	10	600	440	10	6	8,6667	73,333	1,6667	4	0,6
114	2	2	5	20	48	5	2	3	8	2,5	3	0,4
178	2	8	9	144	212	9	2	6,3333	35,333	4,5	7	0,22222
245	3	5	6	90	126	6	3	4,6667	21	2	3	0,5
222	3	3	3	27	54	3	3	3	9	1	0	1
199	2	10	10	200	280	10	2	7,3333	46,667	5	8	0,2
577	6	8	8	384	320	8	6	7,3333	53,333	1,3333	2	0,75
457	5	6	8	240	236	8	5	6,3333	39,333	1,6	3	0,625
389	4	9	10	360	332	10	4	7,6667	55,333	2,5	6	0,4
145	2	5	6	60	104	6	2	4,3333	17,333	3	4	0,33333
026	1	3	7	21	62	7	1	3,6667	10,333	7	6	0,14286

157	2	6	8	96	152	8	2	5,3333	25,333	4	6	0,25
156	2	6	7	84	136	7	2	5	22,667	3,5	5	0,28571
466	5	7	7	245	238	7	5	6,3333	39,667	1,4	2	0,71429
788	8	9	9	648	450	9	8	8,6667	75	1,125	1	0,88889
233	3	4	4	48	80	4	3	3,6667	13,333	1,3333	1	0,75
666	7	7	7	343	294	7	7	7	49	1	0	1
557	6	6	8	288	264	8	6	6,6667	44	1,3333	2	0,75
249	3	5	10	150	190	10	3	6	31,667	3,3333	7	0,3
499	5	10	10	500	400	10	5	8,3333	66,667	2	5	0,5
455	5	6	6	180	192	6	5	5,6667	32	1,2	1	0,83333
007	1	1	8	8	34	8	1	3,3333	5,6667	8	7	0,125
044	1	5	5	25	70	5	1	3,6667	11,667	5	4	0,2
057	1	6	8	48	124	8	1	5	20,667	8	7	0,125
299	3	10	10	300	320	10	3	7,6667	53,333	3,3333	7	0,3
589	6	9	10	540	408	10	6	8,3333	68	1,6667	4	0,6
059	1	6	10	60	152	10	1	5,6667	25,333	10	9	0,1
356	4	6	7	168	188	7	4	5,6667	31,333	1,75	3	0,57143
126	2	3	7	42	82	7	2	4	13,667	3,5	5	0,28571
224	3	3	5	45	78	5	3	3,6667	13	1,6667	2	0,6
347	4	5	8	160	184	8	4	5,6667	30,667	2	4	0,5
068	1	7	9	63	158	9	1	5,6667	26,333	9	8	0,11111
155	2	6	6	72	120	6	2	4,6667	20	3	4	0,33333
679	7	8	10	560	412	10	7	8,3333	68,667	1,4286	3	0,7
444	5	5	5	125	150	5	5	5	25	1	0	1
358	4	6	9	216	228	9	4	6,3333	38	2,25	5	0,44444
045	1	5	6	30	82	6	1	4	13,667	6	5	0,16667
489	5	9	10	450	370	10	5	8	61,667	2	5	0,5
555	6	6	6	216	216	6	6	6	36	1	0	1
269	3	7	10	210	242	10	3	6,6667	40,333	3,3333	7	0,3
049	1	5	10	50	130	10	1	5,3333	21,667	10	9	0,1
139	2	4	10	80	136	10	2	5,3333	22,667	5	8	0,2
035	1	4	6	24	68	6	1	3,6667	11,333	6	5	0,16667
889	9	9	10	810	522	10	9	9,3333	87	1,1111	1	0,9
038	1	4	9	36	98	9	1	4,6667	16,333	9	8	0,11111
015	1	2	6	12	40	6	1	3	6,6667	6	5	0,16667
148	2	5	9	90	146	9	2	5,3333	24,333	4,5	7	0,22222
379	4	8	10	320	304	10	4	7,3333	50,667	2,5	6	0,4
117	2	2	8	32	72	8	2	4	12	4	6	0,25
036	1	4	7	28	78	7	1	4	13	7	6	0,14286
448	5	5	9	225	230	9	5	6,3333	38,333	1,8	4	0,55556
348	4	5	9	180	202	9	4	6	33,667	2,25	5	0,44444
137	2	4	8	64	112	8	2	4,6667	18,667	4	6	0,25
377	4	8	8	256	256	8	4	6,6667	42,667	2	4	0,5
128	2	3	9	54	102	9	2	4,6667	17	4,5	7	0,22222
248	3	5	9	135	174	9	3	5,6667	29	3	6	0,33333
266	3	7	7	147	182	7	3	5,6667	30,333	2,3333	4	0,42857
467	5	7	8	280	262	8	5	6,6667	43,667	1,6	3	0,625
234	3	4	5	60	94	5	3	4	15,667	1,6667	2	0,6
778	8	8	9	576	416	9	8	8,3333	69,333	1,125	1	0,88889
158	2	6	9	108	168	9	2	5,6667	28	4,5	7	0,22222
023	1	3	4	12	38	4	1	2,6667	6,3333	4	3	0,25
469	5	7	10	350	310	10	5	7,3333	51,667	2	5	0,5
167	2	7	8	112	172	8	2	5,6667	28,667	4	6	0,25
478	5	8	9	360	314	9	5	7,3333	52,333	1,8	4	0,55556
013	1	2	4	8	28	4	1	2,3333	4,6667	4	3	0,25

027	1	3	8	24	70	8	1	4	11,667	8	7	0,125
388	4	9	9	324	306	9	4	7,3333	51	2,25	5	0,44444
034	1	4	5	20	58	5	1	3,3333	9,6667	5	4	0,2
447	5	5	8	200	210	8	5	6	35	1,6	3	0,625
479	5	8	10	400	340	10	5	7,6667	56,667	2	5	0,5
244	3	5	5	75	110	5	3	4,3333	18,333	1,6667	2	0,6
278	3	8	9	216	246	9	3	6,6667	41	3	6	0,33333
022	1	3	3	9	30	3	1	2,3333	5	3	2	0,33333
558	6	6	9	324	288	9	6	7	48	1,5	3	0,66667
899	9	10	10	900	560	10	9	9,6667	93,333	1,1111	1	0,9
779	8	8	10	640	448	10	8	8,6667	74,667	1,25	2	0,8
346	4	5	7	140	166	7	4	5,3333	27,667	1,75	3	0,57143
334	4	4	5	80	112	5	4	4,3333	18,667	1,25	1	0,8
014	1	2	5	10	34	5	1	2,6667	5,6667	5	4	0,2
077	1	8	8	64	160	8	1	5,6667	26,667	8	7	0,125
357	4	6	8	192	208	8	4	6	34,667	2	4	0,5
046	1	5	7	35	94	7	1	4,3333	15,667	7	6	0,14286
226	3	3	7	63	102	7	3	4,3333	17	2,3333	4	0,42857
227	3	3	8	72	114	8	3	4,6667	19	2,6667	5	0,375
048	1	5	9	45	118	9	1	5	19,667	9	8	0,11111
458	5	6	9	270	258	9	5	6,6667	43	1,8	4	0,55556
134	2	4	5	40	76	5	2	3,6667	12,667	2,5	3	0,4
016	1	2	7	14	46	7	1	3,3333	7,6667	7	6	0,14286
025	1	3	6	18	54	6	1	3,3333	9	6	5	0,16667
578	6	8	9	432	348	9	6	7,6667	58	1,5	3	0,66667
004	1	1	5	5	22	5	1	2,3333	3,6667	5	4	0,2
236	3	4	7	84	122	7	3	4,6667	20,333	2,3333	4	0,42857
000	1	1	1	1	6	1	1	1	1	1	0	1
677	7	8	8	448	352	8	7	7,6667	58,667	1,1429	1	0,875
168	2	7	9	126	190	9	2	6	31,667	4,5	7	0,22222
189	2	9	10	180	256	10	2	7	42,667	5	8	0,2
777	8	8	8	512	384	8	8	8	64	1	0	1
006	1	1	7	7	30	7	1	3	5	7	6	0,14286
345	4	5	6	120	148	6	4	5	24,667	1,5	2	0,66667
488	5	9	9	405	342	9	5	7,6667	57	1,8	4	0,55556
336	4	4	7	112	144	7	4	5	24	1,75	3	0,57143
116	2	2	7	28	64	7	2	3,6667	10,667	3,5	5	0,28571
378	4	8	9	288	280	9	4	7	46,667	2,25	5	0,44444
339	4	4	10	160	192	10	4	6	32	2,5	6	0,4
146	2	5	7	70	118	7	2	4,6667	19,667	3,5	5	0,28571
135	2	4	6	48	88	6	2	4	14,667	3	4	0,33333
366	4	7	7	196	210	7	4	6	35	1,75	3	0,57143
124	2	3	5	30	62	5	2	3,3333	10,333	2,5	3	0,4
039	1	4	10	40	108	10	1	5	18	10	9	0,1
667	7	7	8	392	322	8	7	7,3333	53,667	1,1429	1	0,875
255	3	6	6	108	144	6	3	5	24	2	3	0,5
099	1	10	10	100	240	10	1	7	40	10	9	0,1
179	2	8	10	160	232	10	2	6,6667	38,667	5	8	0,2
445	5	5	6	150	170	6	5	5,3333	28,333	1,2	1	0,83333
118	2	2	9	36	80	9	2	4,3333	13,333	4,5	7	0,22222
008	1	1	9	9	38	9	1	3,6667	6,3333	9	8	0,11111
349	4	5	10	200	220	10	4	6,3333	36,667	2,5	6	0,4
133	2	4	4	32	64	4	2	3,3333	10,667	2	2	0,5
237	3	4	8	96	136	8	3	5	22,667	2,6667	5	0,375
088	1	9	9	81	198	9	1	6,3333	33	9	8	0,11111

258	3	6	9	162	198	9	3	6	33	3	6	0,33333
449	5	5	10	250	250	10	5	6,6667	41,667	2	5	0,5
125	2	3	6	36	72	6	2	3,6667	12	3	4	0,33333
368	4	7	9	252	254	9	4	6,6667	42,333	2,25	5	0,44444
127	2	3	8	48	92	8	2	4,3333	15,333	4	6	0,25
359	4	6	10	240	248	10	4	6,6667	41,333	2,5	6	0,4
567	6	7	8	336	292	8	6	7	48,667	1,3333	2	0,75
028	1	3	9	27	78	9	1	4,3333	13	9	8	0,11111
239	3	4	10	120	164	10	3	5,6667	27,333	3,3333	7	0,3

ANEXO V

Hiper-Paralelepípedos Inteiros de Dimensão 4

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25
0007	1	1	1	8	8	27	25	8	1	2,75	4,5	6,25	8	8	8	1	1	1	7	7	7	0	0	0	0,125
1469	2	5	7	10	700	199	660	10	2	6	33,167	165	1,4286	2	5	1,4	3,5	2,5	3	5	8	2	5	3	0,2
1169	2	2	7	10	280	142	348	10	2	5,25	23,667	87	1,4286	5	5	3,5	3,5	1	3	8	8	5	5	0	0,2
3458	4	5	6	9	1080	209	786	9	4	6	34,833	196,5	1,5	1,8	2,25	1,2	1,5	1,3	3	4	5	1	2	1	0,4444
2245	3	3	5	6	270	105	279	6	3	4,25	17,5	69,75	1,2	2	2	1,6667	1,6667	1	1	3	3	2	2	0	0,5
1348	2	4	5	9	360	137	382	9	2	5	22,833	95,5	1,8	2,25	4,5	1,25	2,5	2	4	5	7	1	3	2	0,2222
0038	1	1	4	9	36	63	85	9	1	3,75	10,5	21,25	2,25	9	9	4	4	1	5	8	8	3	3	0	0,1111
7778	8	8	8	9	4608	408	2240	9	8	8,25	68	560	1,125	1,125	1,125	1	1	1	1	1	1	0	0	0	0,8889
1689	2	7	9	10	1260	275	1076	10	2	7	45,833	269	1,1111	1,4286	5	1,2857	4,5	3,5	1	3	8	2	7	5	0,2
0056	1	1	6	7	42	69	97	7	1	3,75	11,5	24,25	1,1667	7	7	6	6	1	1	6	6	5	5	0	0,1429
4459	5	5	6	10	1500	245	1000	10	5	6,5	40,833	250	1,6667	2	2	1,2	1,2	1	4	5	5	1	1	0	0,5
0012	1	1	2	3	6	17	17	3	1	1,75	2,8333	4,25	1,5	3	3	2	2	1	1	2	2	1	1	0	0,3333
0689	1	7	9	10	630	249	853	10	1	6,75	41,5	213,25	1,1111	1,4286	10	1,2857	9	7	1	3	9	2	8	6	0,1
5679	6	7	8	10	3360	356	1796	10	6	7,75	59,333	449	1,25	1,4286	1,6667	1,1429	1,3333	1,2	2	3	4	1	2	1	0,6
3334	4	4	4	5	320	108	304	5	4	4,25	18	76	1,25	1,25	1,25	1	1	1	1	1	1	0	0	0	0,8
2247	3	3	5	8	360	127	357	8	3	4,75	21,167	89,25	1,6	2,6667	2,6667	1,6667	1,6667	1	3	5	5	2	2	0	0,375
1359	2	4	6	10	480	164	488	10	2	5,5	27,333	122	1,6667	2,5	5	1,5	3	2	4	6	8	2	4	2	0,2
2236	3	3	4	7	252	103	267	7	3	4,25	17,167	66,75	1,75	2,3333	2,3333	1,3333	1,3333	1	3	4	4	1	1	0	0,4286
2456	3	5	6	7	630	161	531	7	3	5,25	26,833	132,75	1,1667	1,4	2,3333	1,2	2	1,7	1	2	4	1	3	2	0,4286
0089	1	1	9	10	90	129	199	10	1	5,25	21,5	49,75	1,1111	10	10	9	9	1	1	9	9	8	8	0	0,1
2249	3	3	5	10	450	149	435	10	3	5,25	24,833	108,75	2	3,3333	3,3333	1,6667	1,6667	1	5	7	7	2	2	0	0,3
1147	2	2	5	8	160	96	212	8	2	4,25	16	53	1,6	4	4	2,5	2,5	1	3	6	6	3	3	0	0,25
0347	1	4	5	8	160	109	252	8	1	4,5	18,167	63	1,6	2	8	1,25	5	4	3	4	7	1	4	3	0,125
5567	6	6	7	8	2016	272	1212	8	6	6,75	45,333	303	1,1429	1,3333	1,3333	1,1667	1,1667	1	1	2	2	1	1	0	0,75
1678	2	7	8	9	1008	239	886	9	2	6,5	39,833	221,5	1,125	1,2857	4,5	1,1429	4	3,5	1	2	7	1	6	5	0,2222
3478	4	5	8	9	1440	245	988	9	4	6,5	40,833	247	1,125	1,8	2,25	1,6	2	1,3	1	4	5	3	4	1	0,4444
0148	1	2	5	9	90	89	163	9	1	4,25	14,833	40,75	1,8	4,5	9	2,5	5	2	4	7	8	3	4	1	0,1111
4555	5	6	6	6	1080	198	756	6	5	5,75	33	189	1	1	1,2	1	1,2	1,2	0	0	1	0	1	1	0,8333
0279	1	3	8	10	240	155	374	10	1	5,5	25,833	93,5	1,25	3,3333	10	2,6667	8	3	2	7	9	5	7	2	0,1
1457	2	5	6	8	480	156	476	8	2	5,25	26	119	1,3333	1,6	4	1,2	3	2,5	2	3	6	1	4	3	0,25
0058	1	1	6	9	54	85	123	9	1	4,25	14,167	30,75	1,5	9	9	6	6	1	3	8	8	5	5	0	0,1111
5556	6	6	6	7	1512	234	972	7	6	6,25	39	243	1,1667	1,1667	1,1667	1	1	1	1	1	1	0	0	0	0,8571
0069	1	1	7	10	70	105	157	10	1	4,75	17,5	39,25	1,4286	10	10	7	7	1	3	9	9	6	6	0	0,1
1138	2	2	4	9	144	92	196	9	2	4,25	15,333	49	2,25	4,5	4,5	2	2	1	5	7	7	2	2	0	0,2222
0078	1	1	8	9	72	107	161	9	1	4,75	17,833	40,25	1,125	9	9	8	8	1	1	8	8	7	7	0	0,1111
2268	3	3	7	9	567	168	522	9	3	5,5	28	130,5	1,2857	3	3	2,3333	2,3333	1	2	6	6	4	4	0	0,3333
6667	7	7	7	8	2744	315	1519	8	7	7,25	52,5	379,75	1,1429	1,1429	1,1429	1	1	1	1	1	1	0	0	0	0,875
6789	7	8	9	10	5040	431	2414	10	7	8,5	71,833	603,5	1,1111	1,25	1,4286	1,125	1,2857	1,1	1	2	3	1	2	1	0,7
1249	2	3	5	10	300	131	340	10	2	5	21,833	85	2	3,3333	5	1,6667	2,5	1,5	5	7	8	2	3	1	0,2
1126	2	2	3	7	84	65	124	7	2	3,5	10,833	31	2,3333	3,5	3,5	1,5	1,5	1	4	5	5	1	1	0	0,2857
3349	4	4	5	10	800	186	640	10	4	5,75	31	160	2	2,5	2,5	1,25	1,25	1	5	6	6	1	1	0	0,4
3469	4	5	7	10	1400	243	970	10	4	6,5	40,5	242,5	1,4286	2	2,5	1,4	1,75	1,3	3	5	6	2	3	1	0,4
0146	1	2	5	7	70	73	129	7	1	3,75	12,167	32,25	1,4	3,5	7	2,5	5	2	2	5	6	3	4	1	0,1429
3589	4	6	9	10	2160	304	1356	10	4	7,25	50,667	339	1,1111	1,6667	2,5	1,5	2,25	1,5	1	4	6	3	5	2	0,4
0008	1	1	1	9	9	30	28	9	1	3	5	7	9	9	9	1	1	1	8	8	8	0	0	0	0,1111
4445	5	5	5	6	750	165	575	6	5	5,25	27,5	143,75	1,2	1,2	1,2	1	1	1	1	1	1	0	0	0	0,8333

7888	8	9	9	9	5832	459	2673	9	8	8,75	76,5	668,25	1	1	1,125	1	1,125	1,1	0	0	1	0	1	1	0,8889
5559	6	6	6	10	2160	288	1296	10	6	7	48	324	1,6667	1,6667	1,6667	1	1	1	4	4	4	0	0	0	0,6
3578	4	6	8	9	1728	266	1128	9	4	6,75	44,333	282	1,125	1,5	2,25	1,3333	2	1,5	1	3	5	2	4	2	0,4444
2478	3	5	8	9	1080	223	831	9	3	6,25	37,167	207,75	1,125	1,8	3	1,6	2,6667	1,7	1	4	6	3	5	2	0,3333
0059	1	1	6	10	60	93	136	10	1	4,5	15,5	34	1,6667	10	10	6	6	1	4	9	9	5	5	0	0,1
0025	1	1	3	6	18	37	45	6	1	2,75	6,1667	11,25	2	6	6	3	3	1	3	5	5	2	2	0	0,1667
1119	2	2	2	10	80	72	128	10	2	4	12	32	5	5	5	1	1	1	8	8	8	0	0	0	0,2
2228	3	3	3	9	243	108	270	9	3	4,5	18	67,5	3	3	3	1	1	1	6	6	6	0	0	0	0,3333
2238	3	3	4	9	324	123	333	9	3	4,75	20,5	83,25	2,25	3	3	1,3333	1,3333	1	5	6	6	1	1	0	0,3333
2777	3	8	8	8	1536	264	1088	8	3	6,75	44	272	1	1	2,6667	1	2,6667	2,7	0	0	5	0	5	5	0,375
1235	2	3	4	6	144	80	180	6	2	3,75	13,333	45	1,5	2	3	1,3333	2	1,5	2	3	4	1	2	1	0,3333
0013	1	1	2	4	8	21	22	4	1	2	3,5	5,5	2	4	4	2	2	1	2	3	3	1	1	0	0,25
3356	4	4	6	7	672	162	544	7	4	5,25	27	136	1,1667	1,75	1,75	1,5	1,5	1	1	3	3	2	2	0	0,5714
0789	1	8	9	10	720	269	962	10	1	7	44,833	240,5	1,1111	1,25	10	1,125	9	8	1	2	9	1	8	7	0,1
0888	1	9	9	9	729	270	972	9	1	7	45	243	1	1	9	1	9	9	0	0	8	0	8	8	0,1111
0578	1	6	8	9	432	197	606	9	1	6	32,833	151,5	1,125	1,5	9	1,3333	8	6	1	3	8	2	7	5	0,1111
0289	1	3	9	10	270	169	417	10	1	5,75	28,167	104,25	1,1111	3,3333	10	3	9	3	1	7	9	6	8	2	0,1
0777	1	8	8	8	512	216	704	8	1	6,25	36	176	1	1	8	1	8	8	0	0	7	0	7	7	0,125
2589	3	6	9	10	1620	279	1152	10	3	7	46,5	288	1,1111	1,6667	3,3333	1,5	3	2	1	4	7	3	6	3	0,3
3679	4	7	8	10	2240	306	1384	10	4	7,25	51	346	1,25	1,4286	2,5	1,1429	2	1,8	2	3	6	1	4	3	0,4
6669	7	7	7	10	3430	357	1813	10	7	7,75	59,5	453,25	1,4286	1,4286	1,4286	1	1	1	3	3	3	0	0	0	0,7
1168	2	2	7	9	252	131	316	9	2	5	21,833	79	1,2857	4,5	4,5	3,5	3,5	1	2	7	7	5	5	0	0,2222
3333	4	4	4	4	256	96	256	4	4	4	16	64	1	1	1	1	1	1	0	0	0	0	0	0	1
1246	2	3	5	7	210	101	247	7	2	4,25	16,833	61,75	1,4	2,3333	3,5	1,6667	2,5	1,5	2	4	5	2	3	1	0,2857
1258	2	3	6	9	324	135	360	9	2	5	22,5	90	1,5	3	4,5	2	3	1,5	3	6	7	3	4	1	0,2222
0125	1	2	3	6	36	47	72	6	1	3	7,8333	18	2	3	6	1,5	3	2	3	4	5	1	2	1	0,1667
3369	4	4	7	10	1120	222	832	10	4	6,25	37	208	1,4286	2,5	2,5	1,75	1,75	1	3	6	6	3	3	0	0,4
3888	4	9	9	9	2916	351	1701	9	4	7,75	58,5	425,25	1	1	2,25	1	2,25	2,3	0	0	5	0	5	5	0,4444
0068	1	1	7	9	63	96	142	9	1	4,5	16	35,5	1,2857	9	9	7	7	1	2	8	8	6	6	0	0,1111
2225	3	3	3	6	162	81	189	6	3	3,75	13,5	47,25	2	2	2	1	1	1	3	3	3	0	0	0	0,5
2226	3	3	3	7	189	90	216	7	3	4	15	54	2,3333	2,3333	2,3333	1	1	1	4	4	4	0	0	0	0,4286
4789	5	8	9	10	3600	377	1930	10	5	8	62,833	482,5	1,1111	1,25	2	1,125	1,8	1,6	1	2	5	1	4	3	0,5
2267	3	3	7	8	504	155	471	8	3	5,25	25,833	117,75	1,1429	2,6667	2,6667	2,3333	2,3333	1	1	5	5	4	4	0	0,375
1467	2	5	7	8	560	171	542	8	2	5,5	28,5	135,5	1,1429	1,6	4	1,4	3,5	2,5	1	3	6	2	5	3	0,25
0035	1	1	4	6	24	45	58	6	1	3	7,5	14,5	1,5	6	6	4	4	1	2	5	5	3	3	0	0,1667
4666	5	7	7	7	1715	252	1078	7	5	6,5	42	269,5	1	1	1,4	1	1,4	1,4	0	0	2	0	2	2	0,7143
0126	1	2	3	7	42	53	83	7	1	3,25	8,8333	20,75	2,3333	3,5	7	1,5	3	2	4	5	6	1	2	1	0,1429
0001	1	1	1	2	2	9	7	2	1	1,25	1,5	1,75	2	2	2	1	1	1	1	1	1	0	0	0	0,5
1568	2	6	7	9	756	203	696	9	2	6	33,833	174	1,2857	1,5	4,5	1,1667	3,5	3	2	3	7	1	5	4	0,2222
1333	2	4	4	4	128	72	160	4	2	3,5	12	40	1	1	2	1	2	2	0	0	2	0	2	2	0,5
1777	2	8	8	8	1024	240	896	8	2	6,5	40	224	1	1	4	1	4	4	0	0	6	0	6	6	0,25
149	1	2	5	10	100	97	180	10	1	4,5	16,167	45	2	5	10	2,5	5	2	5	8	9	3	4	1	0,1
1349	2	4	5	10	400	148	420	10	2	5,25	24,667	105	2	2,5	5	1,25	2,5	2	5	6	8	1	3	2	0,2
4447	5	5	5	8	1000	195	725	8	5	5,75	32,5	181,25	1,6	1,6	1,6	1	1	1	3	3	3	0	0	0	0,625
0346	1	4	5	7	140	99	223	7	1	4,25	16,5	55,75	1,4	1,75	7	1,25	5	4	2	3	6	1	4	3	0,1429
1119	2	2	2	10	80	72	128	10	2	4	12	32	5	5	5	1	1	1	8	8	8	0	0	0	0,2
2279	3	3	8	10	720	197	642	10	3	6	32,833	160,5	1,25	3,3333	3,3333	2,6667	2,6667	1	2	7	7	5	5	0	0,3
2357	3	4	6	8	576	158	504	8	3	5,25	26,333	126	1,3333	2	2,6667	1,5	2	1,3	2	4	5	2	3	1	0,375
1378	2	4	8	9	576	182	568	9	2	5,75	30,333	142	1,125	2,25	4,5	2	4	2	1	5	7	4	6	2	0,2222
0028	1	1	3	9	27	52	66	9	1	3,5	8,6667	16,5	3	9	9	3	3	1	6	8	8	2	2	0	0,1111
4445	5	5	5	6	750	165	575	6	5	5,25	27,5	143,75	1,2	1,2	1,2	1	1	1	1	1	1	0	0	0	0,8333
0037	1	1	4	8	32	57	76	8	1	3,5	9,5	19	2	8	8	4	4	1	4	7	7	3	3	0	0,125
0367	1	4	7	8	224	135	340	8	1	5	22,5	85	1,1429	2	8	1,75	7	4	1	4	7	3	6	3	0,125
4888	5	9	9	9	3645	378	1944	9	5	8	63	486	1	1	1,8	1	1,8	1,8	0	0	4	0	4	4	0,5556
0157	1	2	6	8	96	92	172	8	1	4,25	15,333	43	1,3333	4	8	3	6	2	2	6	7	4	5	1	0,125

2555	3	6	6	6	648	162	540	6	3	5,25	27	135	1	1	2	1	2	2	0	0	3	0	3	3	0,5
4569	5	6	7	10	2100	287	1280	10	5	7	47,833	320	1,4286	1,6667	2	1,1667	1,4	1,2	3	4	5	1	2	1	0,5
0111	1	2	2	2	8	18	20	2	1	1,75	3	5	1	1	2	1	2	2	0	0	1	0	1	1	0,5
3336	4	4	4	7	448	132	400	7	4	4,75	22	100	1,75	1,75	1,75	1	1	1	3	3	3	0	0	0	0,5714
3359	4	4	6	10	960	204	736	10	4	6	34	184	1,6667	2,5	2,5	1,5	1,5	1	4	6	6	2	2	0	0,4
7779	8	8	8	10	5120	432	2432	10	8	8,5	72	608	1,25	1,25	1,25	1	1	1	2	2	2	0	0	0	0,8
1888	2	9	9	9	1458	297	1215	9	2	7,25	49,5	303,75	1	1	4,5	1	4,5	4,5	0	0	7	0	7	7	0,2222
1149	2	2	5	10	200	114	260	10	2	4,75	19	65	2	5	5	2,5	2,5	1	5	8	8	3	3	0	0,2
1113	2	2	2	4	32	36	56	4	2	2,5	6	14	2	2	2	1	1	1	2	2	2	0	0	0	0,5
2679	3	7	8	10	1680	281	1178	10	3	7	46,833	294,5	1,25	1,4286	3,3333	1,1429	2,6667	2,3	2	3	7	1	5	4	0,3
5569	6	6	7	10	2520	310	1452	10	6	7,25	51,667	363	1,4286	1,6667	1,6667	1,1667	1,1667	1	3	4	4	1	1	0	0,6
5589	6	6	9	10	3240	354	1764	10	6	7,75	59	441	1,1111	1,6667	1,6667	1,5	1,5	1	1	4	4	3	3	0	0,6
3335	4	4	4	6	384	120	352	6	4	4,5	20	88	1,5	1,5	1,5	1	1	1	2	2	2	0	0	0	0,6667
0135	1	2	4	6	48	56	92	6	1	3,25	9,3333	23	1,5	3	6	2	4	2	2	4	5	2	3	1	0,1667
0457	1	5	6	8	240	137	358	8	1	5	22,833	89,5	1,3333	1,6	8	1,2	6	5	2	3	7	1	5	4	0,125
1248	2	3	5	9	270	121	309	9	2	4,75	20,167	77,25	1,8	3	4,5	1,6667	2,5	1,5	4	6	7	2	3	1	0,2222
0356	1	4	6	7	168	111	262	7	1	4,5	18,5	65,5	1,1667	1,75	7	1,5	6	4	1	3	6	2	5	3	0,1429
6689	7	7	9	10	4410	405	2191	10	7	8,25	67,5	547,75	1,1111	1,4286	1,4286	1,2857	1,2857	1	1	3	3	2	2	0	0,7
4579	5	6	8	10	2400	308	1420	10	5	7,25	51,333	355	1,25	1,6667	2	1,3333	1,6	1,2	2	4	5	2	3	1	0,5
0138	1	2	4	9	72	77	134	9	1	4	12,833	33,5	2,25	4,5	9	2	4	2	5	7	8	2	3	1	0,1111
4478	5	5	8	9	1800	267	1145	9	5	6,75	44,5	286,25	1,125	1,8	1,8	1,6	1,6	1	1	4	4	3	3	0	0,5556
0158	1	2	6	9	108	101	192	9	1	4,5	16,833	48	1,5	4,5	9	3	6	2	3	7	8	4	5	1	0,1111
6679	7	7	8	10	3920	381	2002	10	7	8	63,5	500,5	1,25	1,4286	1,4286	1,1429	1,1429	1	2	3	3	1	1	0	0,7
1158	2	2	6	9	216	118	276	9	2	4,75	19,667	69	1,5	4,5	4,5	3	3	1	3	7	7	4	4	0	0,2222
5579	6	6	8	10	2880	332	1608	10	6	7,5	55,333	402	1,25	1,6667	1,6667	1,3333	1,3333	1	2	4	4	2	2	0	0,6
1278	2	3	8	9	432	163	462	9	2	5,5	27,167	115,5	1,125	3	4,5	2,6667	4	1,5	1	6	7	5	6	1	0,2222
3567	4	6	7	8	1344	230	920	8	4	6,25	38,333	230	1,1429	1,3333	2	1,1667	1,75	1,5	1	2	4	1	3	2	0,5
1137	2	2	4	8	128	84	176	8	2	4	14	44	2	4	4	2	2	1	4	6	6	2	2	0	0,25
1111	2	2	2	2	16	24	32	2	2	2	4	8	1	1	1	1	1	1	0	0	0	0	0	0	1
0234	1	3	4	5	60	59	107	5	1	3,25	9,8333	26,75	1,25	1,6667	5	1,3333	4	3	1	2	4	1	3	2	0,2
2333	3	4	4	4	192	84	208	4	3	3,75	14	52	1	1	1,3333	1	1,3333	1,3	0	0	1	0	1	1	0,75
0237	1	3	4	8	96	83	164	8	1	4	13,833	41	2	2,6667	8	1,3333	4	3	4	5	7	1	3	2	0,125
1117	2	2	2	8	64	60	104	8	2	3,5	10	26	4	4	4	1	1	1	6	6	6	0	0	0	0,25
1148	2	2	5	9	180	105	236	9	2	4,5	17,5	59	1,8	4,5	4,5	2,5	2,5	1	4	7	7	3	3	0	0,2222
6678	7	7	8	9	3528	359	1841	9	7	7,75	59,833	460,25	1,125	1,2857	1,2857	1,1429	1,1429	1	1	2	2	1	1	0	0,7778
5557	6	6	6	8	1728	252	1080	8	6	6,5	42	270	1,3333	1,3333	1,3333	1	1	1	2	2	2	0	0	0	0,75
1112	2	2	2	3	24	30	44	3	2	2,25	5	11	1,5	1,5	1,5	1	1	1	1	1	1	0	0	0	0,6667
1145	2	2	5	6	120	78	164	6	2	3,75	13	41	1,2	3	3	2,5	2,5	1	1	4	4	3	3	0	0,3333
1123	2	2	3	4	48	44	76	4	2	2,75	7,3333	19	1,3333	2	2	1,5	1,5	1	1	2	2	1	1	0	0,5
1236	2	3	4	7	168	89	206	7	2	4	14,833	51,5	1,75	2,3333	3,5	1,3333	2	1,5	3	4	5	1	2	1	0,2857
3389	4	4	9	10	1440	258	1024	10	4	6,75	43	256	1,1111	2,5	2,5	2,25	2,25	1	1	6	6	5	5	0	0,4
0026	1	1	3	7	21	42	52	7	1	3	7	13	2,3333	7	7	3	3	1	4	6	6	2	2	0	0,1429
3569	4	6	7	10	1680	264	1108	10	4	6,75	44	277	1,4286	1,6667	2,5	1,1667	1,75	1,5	3	4	6	1	3	2	0,4
1369	2	4	7	10	560	180	556	10	2	5,75	30	139	1,4286	2,5	5	1,75	3,5	2	3	6	8	3	5	2	0,2
2789	3	8	9	10	2160	323	1446	10	3	7,5	53,833	361,5	1,1111	1,25	3,3333	1,125	3	2,7	1	2	7	1	6	5	0,3
1459	2	5	6	10	600	182	580	10	2	5,75	30,333	145	1,6667	2	5	1,2	3	2,5	4	5	8	1	4	3	0,2
5555	6	6	6	6	1296	216	864	6	6	6	36	216	1	1	1	1	1	1	0	0	0	0	0	0	1
0267	1	3	7	8	168	119	269	8	1	4,75	19,833	67,25	1,1429	2,6667	8	2,3333	7	3	1	5	7	4	6	2	0,125
8889	9	9	9	10	7290	513	3159	10	9	9,25	85,5	789,75	1,1111	1,1111	1,1111	1	1	1	1	1	1	0	0	0	0,9
1468	2	5	7	9	630	185	601	9	2	5,75	30,833	150,25	1,2857	1,8	4,5	1,4	3,5	2,5	2	4	7	2	5	3	0,2222
0348	1	4	5	9	180	119	281	9	1	4,75	19,833	70,25	1,8	2,25	9	1,25	5	4	4	5	8	1	4	3	0,1111
6668	7	7	7	9	3087	336	1666	9	7	7,5	56	416,5	1,2857	1,2857	1,2857	1	1	1	2	2	2	0	0	0	0,7778
0458	1	5	6	9	270	149	399	9	1	5,25	24,833	99,75	1,5	1,8	9	1,2	6	5	3	4	8	1	5	4	0,1111
1136	2	2	4	7	112	76	156	7	2	3,75	12,667	39	1,75	3,5	3,5	2	2	1	3	5	5	2	2	0	0,2857
2378	3	4	8	9	864	203	708	9	3	6	33,833	177	1,125	2,25	3	2	2,6667	1,3	1	5	6	4	5	1	0,3333

0579	1	6	8	10	480	212	668	10	1	6,25	35,333	167	1,25	1,6667	10	1,3333	8	6	2	4	9	2	7	5	0,1
3378	4	4	8	9	1152	224	848	9	4	6,25	37,333	212	1,125	2,25	2,25	2	2	1	1	5	5	4	4	0	0,4444
1358	2	4	6	9	432	152	444	9	2	5,25	25,333	111	1,5	2,25	4,5	1,5	3	2	3	5	7	2	4	2	0,2222
4489	5	5	9	10	2250	305	1375	10	5	7,25	50,833	343,75	1,1111	2	2	1,8	1,8	1	1	5	5	4	4	0	0,5
1178	2	2	8	9	288	144	356	9	2	5,25	24	89	1,125	4,5	4,5	4	4	1	1	7	7	6	6	0	0,2222
0489	1	5	9	10	450	209	635	10	1	6,25	34,833	158,75	1,1111	2	10	1,8	9	5	1	5	9	4	8	4	0,1
2257	3	3	6	8	432	141	414	8	3	5	23,5	103,5	1,3333	2,6667	2,6667	2	2	1	2	5	5	3	3	0	0,375
1267	2	3	7	8	336	137	370	8	2	5	22,833	92,5	1,1429	2,6667	4	2,3333	3,5	1,5	1	5	6	4	5	1	0,25
0359	1	4	6	10	240	144	364	10	1	5,25	24	91	1,6667	2,5	10	1,5	6	4	4	6	9	2	5	3	0,1
1256	2	3	6	7	252	113	288	7	2	4,5	18,833	72	1,1667	2,3333	3,5	2	3	1,5	1	4	5	3	4	1	0,2857
1347	2	4	5	8	320	126	344	8	2	4,75	21	86	1,6	2	4	1,25	2,5	2	3	4	6	1	3	2	0,25
1139	2	2	4	10	160	100	216	10	2	4,5	16,667	54	2,5	5	5	2	2	1	6	8	8	2	2	0	0,2
2239	3	3	4	10	360	133	366	10	3	5	22,167	91,5	2,5	3,3333	3,3333	1,3333	1,3333	1	6	7	7	1	1	0	0,3
2223	3	3	3	4	108	63	135	4	3	3,25	10,5	33,75	1,3333	1,3333	1,3333	1	1	1	1	1	1	0	0	0	0,75
2666	3	7	7	7	1029	210	784	7	3	6	35	196	1	1	2,3333	1	2,3333	2,3	0	0	4	0	4	4	0,4286
8889	9	9	9	10	7290	513	3159	10	9	9,25	85,5	789,75	1,1111	1,1111	1,1111	1	1	1	1	1	1	0	0	0	0,9
1367	2	4	7	8	448	154	456	8	2	5,25	25,667	114	1,1429	2	4	1,75	3,5	2	1	4	6	3	5	2	0,25
0189	1	2	9	10	180	149	308	10	1	5,5	24,833	77	1,1111	5	10	4,5	9	2	1	8	9	7	8	1	0,1
1179	2	2	8	10	320	156	392	10	2	5,5	26	98	1,25	5	5	4	4	1	2	8	8	6	6	0	0,2
0134	1	2	4	5	40	49	78	5	1	3	8,1667	19,5	1,25	2,5	5	2	4	2	1	3	4	2	3	1	0,2
0368	1	4	7	9	252	147	379	9	1	5,25	24,5	94,75	1,2857	2,25	9	1,75	7	4	2	5	8	3	6	3	0,1111
2367	3	4	7	8	672	173	572	8	3	5,5	28,833	143	1,1429	2	2,6667	1,75	2,3333	1,3	1	4	5	3	4	1	0,375
1222	2	3	3	3	54	45	81	3	2	2,75	7,5	20,25	1	1	1,5	1	1,5	1,5	0	0	1	0	1	1	0,6667
2237	3	3	4	8	288	113	300	8	3	4,5	18,833	75	2	2,6667	2,6667	1,3333	1,3333	1	4	5	5	1	1	0	0,375
3348	4	4	5	9	720	173	584	9	4	5,5	28,833	146	1,8	2,25	2,25	1,25	1,25	1	4	5	5	1	1	0	0,4444
0014	1	1	2	5	10	25	27	5	1	2,25	4,1667	6,75	2,5	5	5	2	2	1	3	4	4	1	1	0	0,2
4449	5	5	5	10	1250	225	875	10	5	6,25	37,5	218,75	2	2	2	1	1	1	5	5	5	0	0	0	0,5
1117	2	2	2	8	64	60	104	8	2	3,5	10	26	4	4	4	1	1	1	6	6	6	0	0	0	0,25
0046	1	1	5	7	35	60	82	7	1	3,5	10	20,5	1,4	7	7	5	5	1	2	6	6	4	4	0	0,1429
3339	4	4	4	10	640	168	544	10	4	5,5	28	136	2,5	2,5	2,5	1	1	1	6	6	6	0	0	0	0,4
4448	5	5	5	9	1125	210	800	9	5	6	35	200	1,8	1,8	1,8	1	1	1	4	4	4	0	0	0	0,5556
5666	6	7	7	7	2058	273	1225	7	6	6,75	45,5	306,25	1	1	1,1667	1	1,1667	1,2	0	0	1	0	1	1	0,8571
2468	3	5	7	9	945	206	744	9	3	6	34,333	186	1,2857	1,8	3	1,4	2,3333	1,7	2	4	6	2	4	2	0,3333
3555	4	6	6	6	864	180	648	6	4	5,5	30	162	1	1	1,5	1	1,5	1,5	0	0	2	0	2	2	0,6667
4679	5	7	8	10	2800	331	1590	10	5	7,5	55,167	397,5	1,25	1,4286	2	1,1429	1,6	1,4	2	3	5	1	3	2	0,5
0159	1	2	6	10	120	110	212	10	1	4,75	18,333	53	1,6667	5	10	3	6	2	4	8	9	4	5	1	0,1
2689	3	7	9	10	1890	301	1299	10	3	7,25	50,167	324,75	1,1111	1,4286	3,3333	1,2857	3	2,3	1	3	7	2	6	4	0,3
1239	2	3	4	10	240	116	284	10	2	4,75	19,333	71	2,5	3,3333	5	1,3333	2	1,5	6	7	8	1	2	1	0,2
4678	5	7	8	9	2520	311	1459	9	5	7,25	51,833	364,75	1,125	1,2857	1,8	1,1429	1,6	1,4	1	2	4	1	3	2	0,5556
0045	1	1	5	6	30	53	71	6	1	3,25	8,8333	17,75	1,2	6	6	5	5	1	1	5	5	4	4	0	0,1667
1146	2	2	5	7	140	87	188	7	2	4	14,5	47	1,4	3,5	3,5	2,5	2,5	1	2	5	5	3	3	0	0,2857
1234	2	3	4	5	120	71	154	5	2	3,5	11,833	38,5	1,25	1,6667	2,5	1,3333	2	1,5	1	2	3	1	2	1	0,4
1458	2	5	6	9	540	169	528	9	2	5,5	28,167	132	1,5	1,8	4,5	1,2	3	2,5	3	4	7	1	4	3	0,2222
4457	5	5	6	8	1200	213	830	8	5	6	35,5	207,5	1,3333	1,6	1,6	1,2	1,2	1	2	3	3	1	1	0	0,625
2289	3	3	9	10	810	213	711	10	3	6,25	35,5	177,75	1,1111	3,3333	3,3333	3	3	1	1	7	7	6	6	0	0,3
3444	4	5	5	5	500	135	425	5	4	4,75	22,5	106,25	1	1	1,25	1	1,25	1,3	0	0	1	0	1	1	0,8
1114	2	2	2	5	40	42	68	5	2	2,75	7	17	2,5	2,5	2,5	1	1	1	3	3	3	0	0	0	0,4
4689	5	7	9	10	3150	353	1745	10	5	7,75	58,833	436,25	1,1111	1,4286	2	1,2857	1,8	1,4	1	3	5	2	4	2	0,5
1379	2	4	8	10	640	196	624	10	2	6	32,667	156	1,25	2,5	5	2	4	2	2	6	8	4	6	2	0,2
5777	6	8	8	8	3072	336	1664	8	6	7,5	56	416	1	1	1,3333	1	1,3333	1,3	0	0	2	0	2	2	0,75
3367	4	4	7	8	896	192	688	8	4	5,75	32	172	1,1429	2	2	1,75	1,75	1	1	4	4	3	3	0	0,5
0247	1	3	5	8	120	95	199	8	1	4,25	15,833	49,75	1,6	2,6667	8	1,6667	5	3	3	5	7	2	4	2	0,125
2248	3	3	5	9	405	138	396	9	3	5	23	99	1,8	3	3	1,6667	1,6667	1	4	6	6	2	2	0	0,3333
2259	3	3	6	10	540	165	504	10	3	5,5	27,5	126	1,6667	3,3333	3,3333	2	2	1	4	7	7	3	3	0	0,3
3335	4	4	4	6	384	120	352	6	4	4,5	20	88	1,5	1,5	1,5	1	1	1	2	2	2	0	0	0	0,6667

2578	3	6	8	9	1296	243	954	9	3	6,5	40,5	238,5	1,125	1,5	3	1,3333	2,6667	2	1	3	6	2	5	3	0,3333
1666	2	7	7	7	686	189	637	7	2	5,75	31,5	159,25	1	1	3,5	1	3,5	3,5	0	0	5	0	5	5	0,2857
1269	2	3	7	10	420	161	452	10	2	5,5	26,833	113	1,4286	3,3333	5	2,3333	3,5	1,5	3	7	8	4	5	1	0,2
2224	3	3	3	5	135	72	162	5	3	3,5	12	40,5	1,6667	1,6667	1,6667	1	1	1	2	2	2	0	0	0	0,6
0004	1	1	1	5	5	18	16	5	1	2	3	4	5	5	5	1	1	1	4	4	4	0	0	0	0,2
1478	2	5	8	9	720	201	674	9	2	6	33,5	168,5	1,125	1,8	4,5	1,6	4	2,5	1	4	7	3	6	3	0,2222
0568	1	6	7	9	378	181	537	9	1	5,75	30,167	134,25	1,2857	1,5	9	1,1667	7	6	2	3	8	1	6	5	0,1111
0006	1	1	1	7	7	24	22	7	1	2,5	4	5,5	7	7	7	1	1	1	6	6	6	0	0	0	0,1429
1589	2	6	9	10	1080	254	948	10	2	6,75	42,333	237	1,1111	1,6667	5	1,5	4,5	3	1	4	8	3	7	4	0,2

ANEXO VI

Paralelepípedos – Caso Contínuo

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
002	0,1783	0,4417	2,9208	0,2300	3,7788	2,9208	0,1783	1,1802	0,6298	16,3860	2,7425	0,0610
067	0,7857	6,4410	7,2230	36,5530	114,5200	7,2230	0,7857	4,8166	19,0860	9,1931	6,4373	0,1088
134	1,0441	3,9226	4,7116	19,2960	54,9920	4,7116	1,0441	3,2261	9,1654	4,5128	3,6675	0,2216
349	3,2664	4,2582	9,3540	130,1000	168,5900	9,3540	3,2664	5,6262	28,0980	2,8637	6,0876	0,3492
249	2,6482	4,5876	9,9023	120,3000	167,6000	9,9023	2,6482	5,7127	27,9330	3,7393	7,2541	0,2674
559	5,1028	5,4645	9,8714	275,2600	264,4000	9,8714	5,1028	6,8129	44,0660	1,9345	4,7686	0,5169
449	4,1397	4,5551	9,4266	177,7500	201,6400	9,4266	4,1397	6,0405	33,6060	2,2772	5,2870	0,4392
188	1,0885	8,1015	8,1037	71,4650	166,5800	8,1037	1,0885	5,7646	27,7640	7,4445	7,0152	0,1343
014	0,4981	1,6796	4,6732	3,9093	22,0260	4,6732	0,4981	2,2836	3,6711	9,3827	4,1752	0,1066
356	3,7050	5,1851	6,1615	118,3700	147,9700	6,1615	3,7050	5,0172	24,6620	1,6630	2,4565	0,6013
246	2,0308	4,0300	6,8495	56,0580	99,3960	6,8495	2,0308	4,3034	16,5660	3,3727	4,8187	0,2965
569	5,8065	6,0172	9,7960	342,2600	301,5300	9,7960	5,8065	7,2065	50,2540	1,6871	3,9895	0,5927
025	0,5191	2,3144	5,9770	7,1813	36,2750	5,9770	0,5191	2,9368	6,0458	11,5130	5,4578	0,0869
068	0,9071	6,5657	8,2995	49,4310	135,9500	8,2995	0,9071	5,2574	22,6590	9,1494	7,3924	0,1093
005	0,6933	0,8106	5,8148	3,2674	18,6130	5,8148	0,6933	2,4395	3,1021	8,3878	5,1216	0,1192
267	2,3251	6,7873	7,6385	120,5400	170,7700	7,6385	2,3251	5,5836	28,4620	3,2852	5,3134	0,3044
149	1,6356	4,8689	9,1693	73,0210	135,2100	9,1693	1,6356	5,2246	22,5350	5,6061	7,5337	0,1784
089	0,7048	8,3975	9,4960	56,2030	184,7100	9,4960	0,7048	6,1994	30,7850	13,4730	8,7912	0,0742
059	0,6794	5,9247	9,9632	40,1030	139,6500	9,9632	0,6794	5,5224	23,2740	14,6650	9,2838	0,0682
246	2,2797	4,4412	6,6014	66,8350	108,9800	6,6014	2,2797	4,4407	18,1640	2,8958	4,3217	0,3453
129	1,2325	2,0851	9,7460	25,0470	69,8070	9,7460	1,2325	4,3546	11,6350	7,9073	8,5135	0,1265
458	4,0218	5,8714	8,0784	190,7600	207,0700	8,0784	4,0218	5,9905	34,5110	2,0087	4,0566	0,4978
155	1,4624	5,0449	5,1544	38,0270	81,8370	5,1544	1,4624	3,8872	13,6400	3,5247	3,6921	0,2837
579	5,1014	7,6843	9,4972	372,3000	321,2600	9,4972	5,1014	7,4277	53,5430	1,8617	4,3958	0,5372
257	2,2239	5,9353	7,4198	97,9380	147,4800	7,4198	2,2239	5,1930	24,5800	3,3364	5,1959	0,2997
378	3,6774	7,9974	8,7261	256,6300	262,5700	8,7261	3,6774	6,8003	43,7620	2,3729	5,0487	0,4214
011	0,3828	1,1226	1,6459	0,7074	5,8153	1,6459	0,3828	1,0505	0,9692	4,2993	1,2631	0,2326
467	4,7378	6,0874	7,3192	211,1000	216,1500	7,3192	4,7378	6,0482	36,0250	1,5448	2,5814	0,6473
359	3,4663	5,2743	9,5498	174,6000	203,5100	9,5498	3,4663	6,0968	33,9180	2,7550	6,0835	0,3630
235	2,9811	3,3302	5,0725	50,3600	83,8850	5,0725	2,9811	3,7946	13,9810	1,7015	2,0914	0,5877
013	0,4018	1,4065	3,2747	1,8505	12,9730	3,2747	0,4018	1,6943	2,1622	8,1503	2,8729	0,1227
335	3,2273	3,8160	5,5921	68,8700	103,4100	5,5921	3,2273	4,2118	17,2340	1,7327	2,3647	0,5771
157	1,3193	5,0981	7,2860	49,0060	106,9700	7,2860	1,3193	4,5678	17,8280	5,5224	5,9666	0,1811
137	1,4299	3,0467	7,4558	32,4790	75,4640	7,4558	1,4299	3,9774	12,5770	5,2144	6,0259	0,1918
378	3,8195	7,2610	8,1777	226,8000	236,6900	8,1777	3,8195	6,4194	39,4490	2,1410	4,3581	0,4671
678	6,3681	7,1298	8,1140	368,4100	309,8500	8,1140	6,3681	7,2040	51,6420	1,2742	1,7459	0,7848
447	4,1917	4,8641	7,8035	159,1000	182,1100	7,8035	4,1917	5,6198	30,3520	1,8616	3,6117	0,5372
445	4,2787	4,3212	5,7960	107,1600	136,6700	5,7960	4,2787	4,7986	22,7780	1,3546	1,5173	0,7382
112	1,4046	1,7125	2,5517	6,1378	20,7180	2,5517	1,4046	1,8896	3,4531	1,8166	1,1471	0,5505
356	3,6669	5,7542	6,1044	128,8000	157,2200	6,1044	3,6669	5,1752	26,2030	1,6648	2,4376	0,6007
269	2,0306	6,2304	9,7782	123,7100	186,8600	9,7782	2,0306	6,0131	31,1430	4,8153	7,7475	0,2077
449	4,1763	4,5930	9,0273	173,1600	196,6900	9,0273	4,1763	5,9322	32,7820	2,1616	4,8511	0,4626
268	2,0397	6,0668	8,7414	108,1700	166,4700	8,7414	2,0397	5,6159	27,7450	4,2857	6,7017	0,2333
169	1,5209	6,8558	9,9530	103,7800	187,6000	9,9530	1,5209	6,1099	31,2670	6,5442	8,4321	0,1528
489	4,4245	8,3105	9,1852	337,7400	307,4900	9,1852	4,4245	7,3068	51,2480	2,0760	4,7606	0,4817
029	0,0388	2,8437	9,2079	1,0170	53,3060	9,2079	0,0388	4,0302	8,8844	237,0800	9,1691	0,0042
399	3,7204	9,3266	9,8039	340,1900	325,2200	9,8039	3,7204	7,6170	54,2040	2,6351	6,0835	0,3795

245	2,4042	4,6354	5,5678	62,0500	100,6800	5,5678	2,4042	4,2025	16,7800	2,3158	3,1636	0,4318
339	3,3792	3,4449	9,8414	114,5600	157,6000	9,8414	3,3792	5,5552	26,2670	2,9124	6,4623	0,3434
279	2,0945	7,0576	9,0304	133,4900	194,8600	9,0304	2,0945	6,0608	32,4760	4,3115	6,9359	0,2319
237	2,3446	3,1622	7,2881	54,0340	95,0960	7,2881	2,3446	4,2650	15,8490	3,1085	4,9435	0,3217
488	4,0382	8,9489	8,9784	324,4600	305,4800	8,9784	4,0382	7,3218	50,9140	2,2234	4,9402	0,4498
034	0,7064	3,2610	4,3999	10,1360	39,5200	4,3999	0,7064	2,7891	6,5866	6,2286	3,6935	0,1606
239	2,4792	3,9518	9,8324	96,3320	146,0600	9,8324	2,4792	5,4212	24,3430	3,9659	7,3532	0,2522
027	0,1986	2,8037	7,8280	4,3583	48,1170	7,8280	0,1986	3,6101	8,0194	39,4190	7,6294	0,0254
569	5,5459	6,4223	9,9686	355,0500	309,8500	9,9686	5,5459	7,3123	51,6410	1,7975	4,4227	0,5563
112	1,0013	1,9842	2,9107	5,7829	21,3530	2,9107	1,0013	1,9654	3,5589	2,9069	1,9094	0,3440
599	5,3078	9,1886	9,6594	471,1000	377,6000	9,6594	5,3078	8,0519	62,9330	1,8198	4,3515	0,5495
014	0,9788	1,3401	4,0558	5,3196	21,4330	4,0558	0,9788	2,1249	3,5721	4,1438	3,0770	0,2413
224	2,0505	2,0978	4,4150	18,9910	45,2320	4,4150	2,0505	2,8544	7,5387	2,1532	2,3645	0,4644
013	0,2401	1,6415	3,0920	1,2187	12,4240	3,0920	0,2401	1,6579	2,0707	12,8780	2,8519	0,0777
777	7,6461	7,7705	7,8764	467,9700	361,6800	7,8764	7,6461	7,7643	60,2810	1,0301	0,2303	0,9708
457	4,4123	5,6014	7,9449	196,3600	208,5400	7,9449	4,4123	5,9862	34,7570	1,8006	3,5326	0,5554
056	0,9486	5,5715	6,6511	35,1530	97,3030	6,6511	0,9486	4,3904	16,2170	7,0114	5,7025	0,1426
078	0,0338	7,0150	8,7356	2,0684	123,6200	8,7356	0,0338	5,2615	20,6040	258,8100	8,7019	0,0039
026	0,5837	2,1351	6,0983	7,5993	35,6510	6,0983	0,5837	2,9390	5,9419	10,4480	5,5146	0,0957
379	3,0008	7,8775	9,3884	221,9300	251,5400	9,3884	3,0008	6,7556	41,9230	3,1286	6,3876	0,3196
028	0,4760	2,8557	8,1924	11,1350	57,3070	8,1924	0,4760	3,8414	9,5512	17,2120	7,7164	0,0581
179	1,4836	7,6490	9,0832	103,0800	188,6000	9,0832	1,4836	6,0719	31,4340	6,1224	7,5996	0,1633
126	1,4933	2,8705	6,3627	27,2730	64,1040	6,3627	1,4933	3,5755	10,6840	4,2610	4,8695	0,2347
568	5,8335	6,1937	8,6285	311,7600	279,8100	8,6285	5,8335	6,8852	46,6360	1,4791	2,7950	0,6761
236	2,0607	3,8970	6,2320	50,0470	90,3190	6,2320	2,0607	4,0633	15,0530	3,0242	4,1713	0,3307
478	4,4265	7,0551	8,0835	252,4500	248,0800	8,0835	4,4265	6,5217	41,3470	1,8261	3,6570	0,5476
349	3,1786	4,1871	9,9746	132,7500	173,5600	9,9746	3,1786	5,7801	28,9260	3,1381	6,7960	0,3187
389	3,4858	8,2384	9,7891	281,1200	286,9700	9,7891	3,4858	7,1711	47,8290	2,8083	6,3033	0,3561
267	2,5357	6,0167	7,6838	117,2300	161,9400	7,6838	2,5357	5,4121	26,9910	3,0302	5,1481	0,3300
139	1,9724	3,8938	9,2150	70,7740	123,4800	9,2150	1,9724	5,0271	20,5790	4,6718	7,2425	0,2141
057	0,0422	5,7567	7,0349	1,7109	82,0760	7,0349	0,0422	4,2779	13,6790	166,5200	6,9926	0,0060
479	4,2945	7,0623	9,6686	293,2400	280,2600	9,6686	4,2945	7,0084	46,7110	2,2514	5,3741	0,4442
136	1,5062	3,2762	6,7585	33,3500	74,5120	6,7585	1,5062	3,8469	12,4190	4,4872	5,2523	0,2229
077	0,1956	7,0373	7,8529	10,8070	116,3500	7,8529	0,1956	5,0286	19,3920	40,1570	7,6574	0,0249
044	0,7379	4,0792	4,3527	13,1030	47,9560	4,3527	0,7379	3,0566	7,9927	5,8985	3,6148	0,1695
067	0,5252	6,1401	7,1043	22,9110	101,1600	7,1043	0,5252	4,5899	16,8590	13,5260	6,5791	0,0739
259	2,6234	5,2363	9,5076	130,6000	176,9300	9,5076	2,6234	5,7891	29,4880	3,6241	6,8841	0,2759
678	6,4955	7,0726	8,4139	386,5300	320,2000	8,4139	6,4955	7,3273	53,3670	1,2953	1,9184	0,7720
122	1,3809	2,3338	2,9142	9,3919	28,0970	2,9142	1,3809	2,2096	4,6828	2,1103	1,5333	0,4739
123	1,4014	2,2317	3,2398	10,1320	29,7950	3,2398	1,4014	2,2909	4,9659	2,3119	1,8384	0,4326
227	2,2593	2,3759	7,8888	42,3460	83,8680	7,8888	2,2593	4,1747	13,9780	3,4916	5,6295	0,2864
228	2,8323	2,8773	8,7417	71,2400	116,1200	8,7417	2,8323	4,8171	19,3540	3,0864	5,9094	0,3240
788	7,9781	8,2921	8,6073	569,4200	412,4000	8,6073	7,9781	8,2925	68,7330	1,0789	0,6293	0,9269
256	2,1478	5,5251	6,5161	77,3270	123,7300	6,5161	2,1478	4,7297	20,6220	3,0338	4,3682	0,3296
358	3,8541	5,0736	8,7414	170,9300	195,1900	8,7414	3,8541	5,8897	32,5320	2,2681	4,8873	0,4409
699	6,0303	9,3275	9,6368	542,0500	408,5000	9,6368	6,0303	8,3315	68,0830	1,5981	3,6065	0,6258
699	6,6155	9,4656	9,8351	615,8700	441,5600	9,8351	6,6155	8,6387	73,5930	1,4867	3,2197	0,6726
578	5,8662	7,5121	8,8633	390,5800	325,2800	8,8633	5,8662	7,4138	54,2140	1,5109	2,9971	0,6619
569	5,3768	6,1969	9,2239	307,3300	280,1500	9,2239	5,3768	6,9325	46,6910	1,7155	3,8471	0,5829
048	0,5563	4,9919	8,1317	22,5800	95,7860	8,1317	0,5563	4,5600	15,9640	14,6190	7,5754	0,0684
238	2,4250	3,3853	8,3856	68,8410	113,8600	8,3856	2,4250	4,7320	18,9770	3,4579	5,9605	0,2892
239	2,3974	3,6395	9,1945	80,2250	128,4600	9,1945	2,3974	5,0771	21,4110	3,8352	6,7971	0,2607
566	5,3426	6,0602	6,7754	219,3700	219,2700	6,7754	5,3426	6,0594	36,5450	1,2682	1,4327	0,7885
168	1,9708	6,7699	8,5554	114,1500	176,2500	8,5554	1,9708	5,7654	29,3740	4,3410	6,5846	0,2304
122	1,9451	2,0486	2,1996	8,7647	25,5380	2,1996	1,9451	2,0644	4,2564	1,1308	0,2544	0,8843

456	4,7803	5,5038	6,6288	174,4000	188,9600	6,6288	4,7803	5,6376	31,4940	1,3867	1,8484	0,7212
013	0,1838	1,1936	3,8120	0,8364	10,9400	3,8120	0,1838	1,7298	1,8233	20,7380	3,6282	0,0482
148	1,6101	4,2415	8,5372	58,3040	113,5700	8,5372	1,6101	4,7963	18,9290	5,3021	6,9270	0,1886
145	1,6792	4,9745	5,3517	44,7040	87,9240	5,3517	1,6792	4,0018	14,6540	3,1870	3,6725	0,3138
046	0,8843	4,2811	6,1446	23,2610	71,0490	6,1446	0,8843	3,7700	11,8420	6,9489	5,2604	0,1439
456	4,0709	5,8209	6,3658	150,8400	173,3300	6,3658	4,0709	5,4192	28,8880	1,5637	2,2949	0,6395
078	0,2110	7,6937	8,3227	13,5130	134,8200	8,3227	0,2110	5,4091	22,4710	39,4380	8,1116	0,0254
015	0,0472	1,9270	5,4587	0,4968	21,7350	5,4587	0,0472	2,4776	3,6226	115,5800	5,4114	0,0087
456	4,0756	5,5409	6,7994	153,5500	175,9400	6,7994	4,0756	5,4719	29,3230	1,6683	2,7238	0,5994
129	1,1201	2,1248	9,1832	21,8560	64,3570	9,1832	1,1201	4,1427	10,7260	8,1985	8,0631	0,1220
066	0,4619	6,2754	6,9821	20,2400	99,8780	6,9821	0,4619	4,5731	16,6460	15,1150	6,5201	0,0662
788	7,5107	8,3186	8,8086	550,3400	403,8200	8,8086	7,5107	8,2126	67,3040	1,1728	1,2979	0,8527
247	2,4984	4,2238	7,6269	80,4860	123,6500	7,6269	2,4984	4,7830	20,6080	3,0527	5,1285	0,3276
247	2,4915	4,5568	7,6333	86,6630	130,3100	7,6333	2,4915	4,8939	21,7180	3,0638	5,1418	0,3264
679	6,1739	7,0136	9,6421	417,5100	340,9100	9,6421	6,1739	7,6098	56,8180	1,5617	3,4682	0,6403
378	3,4778	7,4271	8,1795	211,2700	230,0500	8,1795	3,4778	6,3615	38,3420	2,3520	4,7018	0,4252
149	1,8778	4,7395	9,9599	88,6400	149,6100	9,9599	1,8778	5,5257	24,9360	5,3041	8,0821	0,1885
189	1,9690	8,6098	9,2794	157,3100	230,2300	9,2794	1,9690	6,6194	38,3720	4,7128	7,3104	0,2122
012	0,1911	1,1160	2,1162	0,4513	5,9587	2,1162	0,1911	1,1411	0,9931	11,0750	1,9251	0,0903
337	3,8123	3,9428	7,6619	115,1700	148,9000	7,6619	3,8123	5,1390	24,8170	2,0098	3,8496	0,4976
689	6,5699	8,5327	9,1398	512,3700	388,1900	9,1398	6,5699	8,0808	64,6980	1,3912	2,5699	0,7188
027	0,2822	2,3086	7,3934	4,8166	39,6130	7,3934	0,2822	3,3281	6,6022	26,2000	7,1112	0,0382
367	3,4706	6,1153	7,6248	161,8200	188,6300	7,6248	3,4706	5,7369	31,4380	2,1970	4,1543	0,4552
256	2,3546	5,2394	6,9893	86,2250	130,8300	6,9893	2,3546	4,8611	21,8050	2,9684	4,6347	0,3369
067	0,1688	6,2293	7,5703	7,9619	98,9750	7,5703	0,1688	4,6561	16,4960	44,8380	7,4014	0,0223
125	1,3945	2,3216	5,5850	18,0810	47,9830	5,5850	1,3945	3,1004	7,9971	4,0052	4,1906	0,2497
446	4,5125	4,6086	6,0621	126,0700	152,1800	6,0621	4,5125	5,0611	25,3630	1,3434	1,5496	0,7444
044	0,3538	4,0917	4,5009	6,5155	42,9120	4,5009	0,3538	2,9821	7,1520	12,7220	4,1471	0,0786
059	0,3041	5,2422	9,0893	14,4890	104,0100	9,0893	0,3041	4,8785	17,3350	29,8910	8,7853	0,0335
349	3,9923	4,0113	9,8917	158,4100	190,3700	9,8917	3,9923	5,9651	31,7280	2,4777	5,8994	0,4036
018	0,4102	1,9073	8,9943	7,0374	43,2550	8,9943	0,4102	3,7706	7,2091	21,9260	8,5841	0,0456
018	0,4764	1,7143	8,1661	6,6687	37,4120	8,1661	0,4764	3,4523	6,2354	17,1430	7,6897	0,0583
014	0,2472	1,1400	4,2255	1,1907	12,2870	4,2255	0,2472	1,8709	2,0478	17,0950	3,9783	0,0585
335	3,0955	3,8579	5,9218	70,7200	106,2400	5,9218	3,0955	4,2917	17,7060	1,9130	2,8263	0,5227
577	5,7022	7,0970	7,1609	289,7900	264,2400	7,1609	5,7022	6,6534	44,0410	1,2558	1,4587	0,7963
156	1,2329	5,8246	6,0526	43,4660	99,7970	6,0526	1,2329	4,3701	16,6330	4,9091	4,8197	0,2037
035	0,8182	3,4134	5,1246	14,3120	48,9560	5,1246	0,8182	3,1187	8,1593	6,2636	4,3065	0,1597
015	0,9619	1,3041	5,5956	7,0191	27,8680	5,5956	0,9619	2,6205	4,6446	5,8171	4,6337	0,1719
336	3,6787	3,9767	6,9359	101,4700	135,4500	6,9359	3,6787	4,8638	22,5760	1,8854	3,2572	0,5304
259	2,7506	5,6006	9,4569	145,6900	188,7700	9,4569	2,7506	5,9361	31,4610	3,4381	6,7062	0,2909
478	4,6215	7,4019	8,7650	299,8300	279,1900	8,7650	4,6215	6,9295	46,5310	1,8966	4,1435	0,5273
699	6,4315	9,8245	9,8845	624,5600	447,7400	9,8845	6,4315	8,7135	74,6230	1,5369	3,4530	0,6507
034	0,3745	3,9137	4,8578	7,1200	44,5940	4,8578	0,3745	3,0487	7,4323	12,9710	4,4833	0,0771
089	0,0039	8,9467	9,2012	0,3212	164,7800	9,2012	0,0039	6,0506	27,4640	2358,0000	9,1973	0,0004
479	4,5751	7,8166	9,6599	345,4500	310,9300	9,6599	4,5751	7,3505	51,8210	2,1114	5,0849	0,4736
136	1,9458	3,4717	6,6164	44,6960	85,2000	6,6164	1,9458	4,0113	14,2000	3,4003	4,6706	0,2941
125	1,5394	2,9589	5,2145	23,7510	56,0220	5,2145	1,5394	3,2376	9,3369	3,3874	3,6751	0,2952
289	2,1779	8,3153	9,8290	178,0000	242,4900	9,8290	2,1779	6,7741	40,4160	4,5131	7,6511	0,2216
058	0,9795	5,5475	8,9967	48,8870	128,3100	8,9967	0,9795	5,1746	21,3850	9,1847	8,0172	0,1089
368	3,7384	6,3020	8,8928	209,5100	225,6900	8,8928	3,7384	6,3111	37,6160	2,3788	5,1544	0,4204
336	3,1828	3,4683	6,4027	70,6760	107,2500	6,4027	3,1828	4,3512	17,8740	2,0117	3,2199	0,4971
034	0,3408	3,0584	4,5165	4,7070	32,7890	4,5165	0,3408	2,6385	5,4648	13,2540	4,1757	0,0754
349	3,1938	4,5211	9,0124	130,1400	167,9400	9,0124	3,1938	5,5758	27,9900	2,8218	5,8186	0,3544
159	1,0660	5,8876	9,6166	60,3530	146,2900	9,6166	1,0660	5,5234	24,3820	9,0215	8,5506	0,1109
267	2,9632	6,6640	7,1331	140,8600	176,8400	7,1331	2,9632	5,5868	29,4730	2,4073	4,1700	0,4154

078	0,8649	7,1349	8,5156	52,5510	148,5900	8,5156	0,8649	5,5051	24,7650	9,8454	7,6506	0,1016
227	2,5503	2,5556	7,2941	47,5400	87,5210	7,2941	2,5503	4,1333	14,5870	2,8601	4,7439	0,3496
116	1,1495	1,4274	6,4278	10,5470	36,4100	6,4278	1,1495	3,0016	6,0683	5,5918	5,2783	0,1788
122	1,0859	2,6819	2,7729	8,0751	26,7190	2,7729	1,0859	2,1802	4,4532	2,5537	1,6870	0,3916
177	1,1592	7,2813	7,7644	65,5330	147,9500	7,7644	1,1592	5,4016	24,6590	6,6984	6,6053	0,1493
378	3,5124	7,0798	8,9220	221,8700	238,7400	8,9220	3,5124	6,5048	39,7910	2,5401	5,4096	0,3937
567	5,0993	6,0292	7,2716	223,5700	223,3300	7,2716	5,0993	6,1334	37,2220	1,4260	2,1723	0,7013
035	0,3942	3,8185	5,5309	8,3252	49,6110	5,5309	0,3942	3,2479	8,2684	14,0310	5,1367	0,0713
788	7,0475	8,4568	8,8087	524,9900	392,3400	8,8087	7,0475	8,1043	65,3910	1,2499	1,7612	0,8001
238	2,2974	3,5851	8,5626	70,5260	117,2100	8,5626	2,2974	4,8150	19,5350	3,7271	6,2652	0,2683
045	0,1465	4,9506	5,9954	4,3492	62,5700	5,9954	0,1465	3,6975	10,4280	40,9160	5,8489	0,0244
014	0,3006	1,6510	4,8225	2,3936	19,8160	4,8225	0,3006	2,2580	3,3026	16,0400	4,5218	0,0623
068	0,1201	6,4795	8,6508	6,7313	115,7400	8,6508	0,1201	5,0834	19,2900	72,0360	8,5307	0,0139
245	2,7879	4,0044	5,5795	62,2890	98,1230	5,5795	2,7879	4,1239	16,3540	2,0013	2,7916	0,4997
057	0,3460	5,7413	7,6751	15,2480	97,4160	7,6751	0,3460	4,5875	16,2360	22,1800	7,3291	0,0451
138	1,6726	3,8386	8,2792	53,1550	104,1000	8,2792	1,6726	4,5968	17,3500	4,9500	6,6066	0,2020
178	1,7044	7,0457	8,9141	107,0500	180,0200	8,9141	1,7044	5,8881	30,0030	5,2301	7,2097	0,1912
034	0,9654	3,4186	4,0779	13,4590	42,3560	4,0779	0,9654	2,8207	7,0594	4,2239	3,1125	0,2368
347	3,7683	4,0203	7,5593	114,5200	148,0500	7,5593	3,7683	5,1159	24,6750	2,0060	3,7910	0,4985
025	0,8845	2,4338	5,3890	11,6010	40,0710	5,3890	0,8845	2,9025	6,6784	6,0928	4,5045	0,1641
179	1,8974	7,6279	9,8258	142,2100	216,1300	9,8258	1,8974	6,4504	36,0220	5,1784	7,9283	0,1931
459	4,3469	5,3883	9,2170	215,8800	226,3000	9,2170	4,3469	6,3174	37,7170	2,1203	4,8700	0,4716
449	4,0868	4,6190	9,4783	178,9200	202,7900	9,4783	4,0868	6,0614	33,7980	2,3192	5,3915	0,4312
259	2,0564	5,6147	9,9048	114,3600	175,0500	9,9048	2,0564	5,8586	29,1750	4,8166	7,8484	0,2076
024	0,4573	2,9339	4,9339	6,6199	36,1470	4,9339	0,4573	2,7750	6,0246	10,7890	4,4766	0,0927
299	2,5612	9,8376	9,9398	250,4500	296,8800	9,9398	2,5612	7,4462	49,4790	3,8809	7,3785	0,2577
033	0,2568	3,1261	3,1830	2,5548	23,1410	3,1830	0,2568	2,1886	3,8568	12,3970	2,9262	0,0807
245	2,5952	4,5324	5,4868	64,5370	101,7400	5,4868	2,5952	4,2048	16,9560	2,1142	2,8916	0,4730
378	3,2716	7,1667	8,3082	194,8000	220,3400	8,3082	3,2716	6,2488	36,7240	2,5395	5,0366	0,3938
129	1,2288	2,9750	9,3349	34,1240	85,7940	9,3349	1,2288	4,5129	14,2990	7,5970	8,1061	0,1316
578	5,5035	7,4593	8,7716	360,1000	309,5200	8,7716	5,5035	7,2448	51,5860	1,5938	3,2681	0,6274
167	1,1762	6,2936	7,5489	55,8830	127,5800	7,5489	1,1762	5,0062	21,2640	6,4178	6,3726	0,1558
135	1,9659	3,2524	5,2441	33,5310	67,5190	5,2441	1,9659	3,4875	11,2530	2,6675	3,2781	0,3749
245	2,5812	4,6377	5,6260	67,3470	105,1700	5,6260	2,5812	4,2816	17,5280	2,1796	3,0448	0,4588
127	1,8892	2,6595	7,3413	36,8860	76,8360	7,3413	1,8892	3,9633	12,8060	3,8859	5,4521	0,2573
034	0,0362	3,1068	4,3744	0,4922	27,7230	4,3744	0,0362	2,5058	4,6205	120,7800	4,3382	0,0083
139	1,1576	3,1246	9,6023	34,7310	89,4710	9,6023	1,1576	4,6282	14,9120	8,2952	8,4447	0,1206
599	5,5892	9,0366	9,3933	474,4200	375,7800	9,3933	5,5892	8,0063	62,6300	1,6806	3,8041	0,5950
268	2,1730	6,9508	8,9227	134,7700	193,0300	8,9227	2,1730	6,0155	32,1710	4,1061	6,7497	0,2435
789	7,5679	8,0567	9,0121	549,4900	403,5700	9,0121	7,5679	8,2122	67,2610	1,1908	1,4441	0,8398
226	2,6664	2,6795	6,9988	50,0050	89,1200	6,9988	2,6664	4,1149	14,8530	2,6248	4,3324	0,3810
268	2,3378	6,1122	8,0905	115,6100	165,3100	8,0905	2,3378	5,5135	27,5510	3,4608	5,7528	0,2890
145	1,7228	4,1288	5,2950	37,6650	76,1960	5,2950	1,7228	3,7156	12,6990	3,0734	3,5722	0,3254
168	1,5188	6,2334	8,0476	76,1880	143,7100	8,0476	1,5188	5,2666	23,9510	5,2987	6,5289	0,1887
566	5,4796	6,8394	6,9982	262,2700	247,3800	6,9982	5,4796	6,4391	41,2290	1,2771	1,5185	0,7830
025	0,0041	2,1795	5,1467	0,0461	22,4950	5,1467	0,0041	2,4434	3,7491	1253,0000	5,1426	0,0008
027	0,2136	2,8508	7,6368	4,6503	48,0220	7,6368	0,2136	3,5671	8,0037	35,7530	7,4232	0,0280
014	0,6695	1,7079	4,2811	4,8949	22,6420	4,2811	0,6695	2,2195	3,7737	6,3948	3,6116	0,1564

ANEXO VII

Classificações Alunos

	<i>Português</i>	<i>Edu. Física</i>	<i>Inglês</i>	<i>Matemática</i>	<i>Ciências Natureza</i>	<i>Geografia</i>	<i>História</i>	<i>Francês</i>	<i>Edu. Visual</i>
555555555	5	5	5	5	5	5	5	5	5
555555555	5	5	5	5	5	5	5	5	5
444445444	4	4	4	4	4	5	4	4	4
445455444	4	4	5	4	5	5	4	4	4
444455544	4	4	4	4	5	5	5	4	4
243333334	2	4	3	3	3	3	3	3	4
232223334	2	3	2	2	2	3	3	3	4
253223334	2	5	3	2	2	3	3	3	4
353345444	3	5	3	3	4	5	4	4	4
343335334	3	4	3	3	3	5	3	3	4
555555555	5	5	5	5	5	5	5	5	5
355334234	3	5	5	3	3	4	2	3	4
444455545	4	4	4	4	5	5	5	4	5
555555555	5	5	5	5	5	5	5	5	5
232222233	2	3	2	2	2	2	2	3	3
443335445	4	4	3	3	3	5	4	4	5
242324333	2	4	2	3	2	4	3	3	3
243223333	2	4	3	2	2	3	3	3	3
344443344	3	4	4	4	4	3	3	4	4
555555555	5	5	5	5	5	5	5	5	5
454234434	4	5	4	2	3	4	4	3	4
232223233	2	3	2	2	2	3	2	3	3
454344455	4	5	4	3	4	4	4	5	5
445545453	4	4	5	5	4	5	4	5	3
455455554	4	5	5	4	5	5	5	5	4
253323233	2	5	3	3	2	3	2	3	3
455455555	4	5	5	4	5	5	5	5	5
444445455	4	4	4	4	4	5	4	5	5
445444455	4	4	5	4	4	4	4	5	5
243223335	2	4	3	2	2	3	3	3	5
343345444	3	4	3	3	4	5	4	4	4
343333334	3	4	3	3	3	3	3	3	4
242223323	2	4	2	2	2	3	3	2	3
332323334	3	3	2	3	2	3	3	3	4
343234344	3	4	3	2	3	4	3	4	4
444445445	4	4	4	4	4	5	4	4	5
333234333	3	3	3	2	3	4	3	3	3
344355344	3	4	4	3	5	5	3	4	4
443445445	4	4	3	4	4	5	4	4	5
453335345	4	5	3	3	3	5	3	4	5
345445455	3	4	5	4	4	5	4	5	5
232323334	2	3	2	3	2	3	3	3	4
233223333	2	3	3	2	2	3	3	3	3
344445445	3	4	4	4	4	5	4	4	5
333233333	3	3	3	2	3	3	3	3	3
233333334	2	3	3	3	3	3	3	3	4
354444445	3	5	4	4	4	4	4	4	5
244333334	2	4	4	3	3	3	3	3	4
433434345	4	3	3	4	3	4	3	4	5
343223345	3	4	3	2	2	3	3	4	5

