



OTIMIZAÇÃO DE PORTFÓLIOS COM MÉTODO DOS POLINÔMIO CANÔNICOS CRUZADO E *CLUSTERS*

Pedro José Papandréa^{*a}, Anderson Paulo de Paiva^a, Xueping Li^b, Rafael Coradi Leme^a, João Éderson Corrêa^a, Vinicius de Carvalho Paes^a.

^a Universidade Federal de Itajubá, Instituto de Engenharia de Produção e Gestão, UNIFEI, Itajubá – MG

^b University of Tennessee, Department of Industrial and Systems Engineering, Knoxville, Tennessee, USA

RESUMO

A metodologia de engenharia chamada de polinômios canônicos ou projeto de experimentos de misturas, e projeto de misturas cruzado foi aplicada para otimizar portfólios de ativos financeiros. Aqui é apresentado também o método de aglutinação aplicado ao conceito de portfólio, a formação de clusters. Este trabalho foi motivado pela possibilidade de diminuir o número de execuções de experimentos com um grande número de variáveis, o que é aplicado em um novo tipo de estudo. Vinte e quatro diferentes ativos foram usados para as demonstrações e exemplos. Os resultados demonstraram é possível usar o experimento cruzado de mistura para agrupamentos de empresas e pode ser utilizado em qualquer quantidade de ativos. A portfólio ideal foi encontrada usando a função desirability como método de otimização. No final, novas propostas são mostradas para os estudos futuros. A comparação foi feita contra o modelo de portfólio com misturas simples sem clusters.

PALAVRAS-CHAVE:

otimização de portfólios,
experimento de mistura cruzado,
ativos financeiros.

INTRODUÇÃO

Métodos de seleção de portfólios são amplamente estudados e desenvolvidos. Há diferentes formas de fazer essa seleção. Iniciando por Markowitz (1952) e passando por outros trabalhos depois dele, (Markowitz, 1959; Cass & Stiglitz, 1970; Amihud & Barnea, 1974; Mayers, 1974; Li & Ng, 2000; Costa & Paiva, 2002; Bertland & Asgharian, 2008; Chiarawongse *et al.*, 2012), o que corrobora com a grande relevância deste tema. Da mesma maneira, projetos de experimentos (DOE) e em especial, método dos polinômios canônicos (MDE) têm sido usados para otimização de produtos e processos ao longo dos anos (Scheffé, 1958, 1963; Chan, 1988; Duineveld *et al.*, 1993; Eriksson *et al.*, 1998; Vieira & Dal Bello, 2006; Zhang & Wong, 2013). MDE é comumente aplicado nos campos de estudo da química e indústrias de alimentos. De Oliveira *et al.* (2011), utilizaram pela primeira vez estes modelos de otimização em mercados de energia elétrica.

Há uma importante restrição no MDE, o número de componentes na mistura. Quando o número de componentes aumenta, a quantidade de experimentos aumenta exponencialmente, sendo n^{k-1} na qual n é o número de experimentos ou pontos e k a quantidade de componentes. Isto torna claro que quando k aumenta, o número de experimentos eleva-se exponencialmente. Scheffé (1963) descreveu isso como q componentes compreendem $2^q - 1$ pontos representando as misturas, sendo que consistem em todas as possibilidades de subconjuntos com q componentes, presentes em iguais proporções. Quando a simulação é utilizada na condução de experimentos, o número máximo de componentes é limitado pelos recursos computacionais disponíveis, como memória, processadores. A proposta é aplicar experimentos cruzados de mistura (Piepel, 1999) para resolver também esta questão e melhorar a robustez das

* Autor para correspondência:
Pedro José Papandrea
e-mail: pedro.papandrea@gmail.com

funções do portfólio. O estudo de caso se utiliza como objeto de ativos da Nasdaq-EUA e BMF&Bovespa-Brasil para demonstrar a eficiência da proposta.

MATERIAIS E MÉTODOS

Literatura

Esta sessão tem como objetivo apresentar conceitos que darão embasamento ao trabalho apresentado. Iniciando projeto de experimentos de mistura cruzado (CM-DOE), passando por projeto de experimentos (DOE), projeto de experimentos de misturas (MDE), portfólio e seleção de portfólio por polinômios canônicos.

Projeto de experimentos de mistura cruzado

Experimentos são os métodos mais eficientes para novos descobrimentos, eles são um de várias estratégias de pesquisa. São baseadas em inferências sobre o meio ambiente para tornar possível o estudo de como objetos de estudo reagem a diferentes estímulos, descobrindo assim como eles funcionam (Kennedy & Bush, 1985). Há várias alternativas para se proceder, e.g., observação passiva – quando o cientista assiste o processo esperando que algo interessante aconteça. Este método oferece grandes massas de dados, mas esses dados são difíceis de identificar qual informação é realmente relevante para o objeto problema da pesquisa. Por outro lado, métodos de experimentação adaptam os dados a problemas específicos. Isto significa que experimentos demandam menos esforços e provêm menos dados irrelevantes (Andersson, 2012). A grande diferença é que o observador investiga apenas estados preexistentes enquanto que o experimentador altera o estado e investiga os resultados das alterações.

CM-DOE foi introduzido pela combinação de dois tipos diferentes de misturas para criar uma região de interesse como uma terceira mistura (Cornell, 1971) e depois como experimento de mistura de dois estágios (Cornell & Ramsey, 1997). Este é também conhecido como projeto de mistura de mistura (MoM) (Piepel, 1999; Borges *et al.*, 2007; Di Zio *et al.*, 2007; Kang *et al.*, 2011). É um tipo de experimento usado para otimizar produtos, principalmente quando há muitas restrições nas proporções dos componentes das misturas (Kang *et al.*, 2011).

Cornell & Ramsey (1997) dividiram os componentes em duas categorias, os componentes principais e os menores. Os componentes principais são o primeiro estágio, mais elevado da mistura e os componentes menores são as outras fases mais baixas ou subgrupos. Quando pensamos na inclusão de variáveis de processo no projeto de mistura é possível visualizar algumas equações mostradas em Cornell (1971) define a função linear para os componentes da mistura com variáveis de processo como sendo w_i ($1 \leq i \leq k - q$). Os w_i 's correspondem aos x_i em uma categoria particular contendo n_i componentes definidos como (1).

$$w_i = \frac{v_{S_l+i+1} \left(\sum_{j=1}^i h_{S_l}^2 + j \right) - \left[\sum_{j=1}^i v_{S_l+j} h_{S_l+j} \right] h_{S_l+i+1}}{\left[\left(\sum_{j=1}^i h_{S_l}^2 + j \right) \left(\sum_{j=1}^{i+1} h_{S_l}^2 + j \right) \right]^{0.5}} \quad (i = 1, \dots, n_l - 1) \quad (1)$$

Há um modelo de DOE que combina polinômios canônicos com variáveis de processo. Cornell (2002) descreveu este tipo de modelo como um arranjo *simplex-lattice* com fatorial. No caso dele, havia um polinômio com q componentes mais n variáveis de processo, sendo esta segunda qualquer número inteiro positivo n .

Método dos polinômios canônicos

O método dos polinômios canônicos ou experimentos de mistura são aqueles cujos fatores são componentes ou ingredientes de uma mistura (Draper & Pukelsheim, 1999; Gozálviz-Zafrilla *et al.*, 2013), i.e., diferentes proporções de qualquer componente podem modificar as características do produto, ou objeto de análise. Eles dependem das proporções dos componentes e não do total da mistura (Scheffé, 1958). Isto significa que as proporções de cada componente ou de cada ingrediente são mais importantes do que a quantidade final do objeto em estudo.

O projeto de misturas possui uma restrição particular que precise ser considerada. A soma das proporções precisa ser igual a um.

Em outras palavras, a soma deve ser igual a cem por cento. Isto distingue-o dos outros métodos porque os fatores controláveis, aqueles que podem ser modificados a qualquer tempo pelo pesquisador, representam as proporções dos componentes da mistura. As proporções, neste caso, são não negativas (Cornell, 2002). Em caso de investimentos do tipo *short sale* (venda a descoberto), o modelo pode aceitar proporções negativas.

Considerando cada proporção da mistura como um fator peso w_i para q componentes, tem-se (2).

$$w_i \geq 0, \quad i = 1, 2, \dots, q \quad (2a)$$

$$\sum_{i=1}^q w_i = w_1 + w_2 + \dots + w_q = 1.0 \quad (2b)$$

$$\sum_{i=1}^q w_i = 1.0 \quad (2c)$$

Um importante método para resolver o problema de otimização de misturas é a superfície de resposta. Ele assume que existe uma função de relacionamento entre os componentes e que podem ser descritos graficamente como uma superfície (Cornell, 2011a), veja em (3) que η denota o valor da resposta e que é dependente de x_1, x_2, \dots, x_q .

$$\eta = \varphi(x_1, x_2, \dots, x_q) \quad (3)$$

Aplicando a aproximação por série de Taylor no polinômio de grau um, tem-se (4) (Wolfinger & Lin, 1997).

$$\eta = \beta_0 + \sum_{i=1}^q \beta_i x_i \quad (4a)$$

E no polinômio de grau dois:

$$\eta = \beta_0 + \sum_{i=1}^q \beta_i x_i + \sum_{i=1}^q \sum_{i \leq j}^q \beta_{ij} x_i x_j \quad (4b)$$

A área no interior do desenho pode ser restringida bem. Acontece possível usar restrições para cada componente, conforme necessário. O ponto central é que a proporção é um dividido por o número de componentes $w_i = 1/n$.

Há um grande número de aplicações de MDE. Alguns dos que podem ser listadas começando Scheffé (1958), passando por Scheffé (1963) usando baricentro ou ponto central. Brandvik & Daling (1998) em uma otimização no desenvolvimento de produtos. Mage & Naes (2005) usando MDE com variáveis de processo ou projeto fatorial.

Há também vários trabalhos utilizando o conceito original de MDE, i.e., usando MDE a análise e otimização das proporções dos ingredientes num produto (Mazaheri Tehrani *et al.*, 2012; Lonni *et al.*, 2012; Ngun *et al.*, 2014) com ou sem restrições.

Seleção de portfólio

Harry Markowitz introduziu o modelo chamado de seleção de portfólio (Markowitz, 1952). A otimização de seleção de portfólios ou de portfólio é uma ferramenta econométrica para ajudar o investidor a decidir o quanto de sua riqueza colocar em cada investimento. No nosso caso, o tipo de investimento é o trunfo do mercado Stock Exchange. Ele também pode ser aplicado em diferentes tipos de investimentos (Amihud & Barnea, 1974; Mayers, 1974; Trippi, 1989; Better & Glover, 2006; Delarue *et al.*, 2011).

Quando chamamos portfólio, é provável que se lembrar de um cesto cheio de ovos. Se a cesta cai, perdemos todos os ovos. Mas dividindo a mesma quantidade de ovos em cestas diferentes, o risco de perder diminuirá. A isto chamamos de diversificação. O risco pode ser definido em termos de alterações nos valores entre duas datas (Artzner *et al.*, 1999). Pode considerar-se na maioria dos casos, a variância dos retornos em um período conhecido. O investidor sempre quer mais retorno do que menos retorno, menos

risco do que mais riscos. A oportunidade na seleção portfólio é equilibrar essas duas metas inversos, retorno e risco. Markowitz (1952) provou que a escolha de um nível de retorno, o risco do investimento usando a teoria do portfólio pode ser inferior a escolher apenas um investimento. Ele também lembrar-nos sobre um fator importante para a decisão marca, o *know-how* do investidor. Se um dos investidores diversifica entre duas portfólios, a variância ou o risco de o composto resultado será menor do que a variação de qualquer portfólio original. Se, e somente se, os dois portfólios são perfeitamente correlacionados, a variância será não diminui, mas em qualquer outro caso, será aumentada.

Usando seu trabalho, vamos explicar matematicamente como a portfólio é usada. Suponha que exista r ativos com $r_i, (i = 1, 2, \dots, n)$ o retorno de cada um separadamente e r_p é o retorno do portfólio. Como o portfólio é composto por vários ativos, considere então w_i como o peso de cada um no portfólio. Isto pode ser resumido em (5) e o risco em termos de variância em (6), na qual Σ é a matriz de covariância com sua diagonal principal composta por a variação de cada ativo. Com estas equações é possível calcular a fronteira eficiente. É uma fronteira de Pareto que significa que cada ponto ao longo da fronteira é o portfólio ideal com retorno e risco relacionados.

$$\begin{aligned} r_p &= w_i \cdot x_i \\ r_p &= \mathbf{w}^T \cdot \mathbf{r} \end{aligned} \quad (5)$$

$$\sigma_p^2 = \mathbf{w}^T \Sigma \mathbf{w} \quad (6)$$

Existem algumas condições que devem ser satisfeitas para virar prática da fronteira eficiente (7) (Markowitz, 1959).

$$E = \sum_{i=1}^n X_i \mu_i \quad (7a)$$

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n X_i X_j \sigma_{ij} \quad (7b)$$

$$\sum_{i=1}^n X_i = 1.0 \quad (7c)$$

$$X_i \geq 0 \quad \text{for } i = 1, 2, \dots, q \quad (7d)$$

Note que o X_i é o peso aplicado em cada ativo no portfólio. Este pode ser relacionado agora a (2c) a qual w_i é o peso. Este é o primeiro relacionamento entre MDE e portfólio.

Para resolver este problema, a otimização pode ser feita de várias formas, e.g., a linha isomédia (Markowitz, 1952), média-Gini de Shalit & Yitzhaki (1984), método analítico de Li & Ng (2000), regra de minimax de Yu *et al.* (2005), função *desirability* De Oliveira *et al.* (2011), algoritmos de inteligência híbrida de Zhang *et al.* (2012) entre outros. O método de otimização é uma escolha do analista, mas em todos os casos, o objetivo é o mesmo, com um rendimento máximo com a menor risco com ele relacionado.

Polinômios canônicos na seleção de portfólios

Após as explicações anteriores, podemos começar a melhorar o conhecimento sobre MDE aplicado no portfólio e depois de que a partir do conteúdo da proposta. A primeira aplicação do MDE para seleção de portfólios estava com (De Oliveira *et al.*, 2011). Depois disso, este método ainda um campo inexplorado. Lembrando o (2) e (7). É claro que, eles estão relacionados, ou podemos considerá-lo como o mesmo. Podemos usar as propriedades do MDE para gerar as matrizes para a otimização de portfólios. Esta metodologia tem as abordagens: pequeno esforço computacional; simples análise de resposta; a possibilidade de utilizar a função *desirability* multi-objetivo; inclusão de aversão ao risco no modelo; possibilidade de inclusão de variáveis de processo no portfólio. Usando um exemplo de dois ativos com os rendimentos e os riscos apresentados na Tabela 1, a concepção mistura com 0,2 resolução e considerando um coeficiente de correlação de 0,5 (que a resolução pode ser aumentada usando, por exemplo 0,01) é apresentada na Tabela 2. Neste caso, usando a (5) e (6), as matrizes de retorno do portfólio e risco serão apresentadas na Tabela 2.

Tabela 1 - Exemplo dois ativos

	Ativo 1	Ativo 2
μ	0.1507	0.1589
σ^2	0.0367	0.0514
σ	0.1916	0.2242

Tabela 2 - Mistura, retornos esperados e os riscos do exemplo portfólio

w_1	w_2	Retorno	Risco
0.000	1.000	0.159	0.227
0.200	0.800	0.157	0.203
0.400	0.600	0.156	0.187
0.600	0.400	0.154	0.179
0.800	0.200	0.152	0.180
1.000	0.000	0.151	0.192

Acontece possível entender o conceito de MDE aplicado a seleção de portfólios. Usando a função de conveniência é possível para encontrar a melhor relação entre os objetivos: o máximo de retorno e risco mínimo.

Arranjo de mistura cruzado na otimização de portfólios

A pesquisa foi realizada de maneira exploratória para a coleta de dados e estudo de caso aplicando o metodolo-gia proposta e analisar os resultados. Como visto, a correlação entre os assuntos ativos. É importante entender como os critérios para a escolha de cada ativo de cada portfólio menor antes começar a proposta método. Considerando-se um exemplo de uma portfólio de três ativos cada, a mistura pode ser escrita como na Tabela 3. Agora, como exemplo, podemos descrever a matriz e as equações de três portfólios de três ativos de cada um. Cada componente da mistura será de outra mistura com os componentes secundários γ_n (8).

Tabela 3 - Mistura um portfólio, três ativos

w_1	w_2	w_3	E	σ_p^2
1.0	0.0	0.0	R_{p1}	S_{p1}
0.0	1.0	0.0	R_{p2}	S_{p2}
0.0	0.0	1.0	R_{p3}	S_{p3}
0.5	0.5	0.0	R_{p4}	S_{p4}
0.0	0.5	0.5	R_{p5}	S_{p5}
0.5	0.0	0.5	R_{p6}	S_{p6}

$$\gamma_1 = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \\ 0.5 & 0.5 & 0.0 \\ 0.0 & 0.5 & 0.5 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} \quad \gamma_2 = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \\ 0.5 & 0.5 & 0.0 \\ 0.0 & 0.5 & 0.5 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} \quad \gamma_3 = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \\ 0.5 & 0.5 & 0.0 \\ 0.0 & 0.5 & 0.5 \\ 0.5 & 0.0 & 0.5 \end{bmatrix} \quad (8)$$

A nova matriz para os principais componentes da mistura será Tabela 4. O retorno (média) e risco (variância) do portfólio cruzado pode ser estimada utilizando a (9).

$$E_\theta = \sum_{i=1}^n w_i \cdot \gamma_i^T \quad (9a)$$

$$\sigma_p^2 = \sum_{i=1}^n w_i^2 (\gamma_i^T)^2 \sigma_{w\gamma} \quad (9b)$$

$$\sum_{i=1}^n w_i = \rho = \text{número de portfólios ou clusters} \quad (9c)$$

$$w_i \geq 0 \quad \text{for } i = 1, 2, \dots, q \quad (9d)$$

Tabela 4 - Mistura três portfólios, três ativos

w_1	w_2	w_3	E	σ_p^2
γ_{11}	γ_{12}	γ_{13}	θ_1	ϕ_1
γ_{21}	γ_{22}	γ_{23}	θ_2	ϕ_2
\vdots	\vdots	\vdots	\vdots	\vdots
γ_{n1}	γ_{n2}	γ_{n3}	θ_n	ϕ_n

A correlação no portfólio

A correlação r_{xy} (Newbold *et al.*, 2012) descreve a relação entre duas variáveis aleatórias, (10). Ele será dimensionado de -1,0, passando para 0,0, para +1,0. Com um coeficiente de correlação de +1,0 essa variável X vontade tem a relação linear positiva perfeita para outro Y . O coeficiente de correlação de 0,0 indica que as duas variáveis não têm qualquer relação.

Finalmente, o coeficiente de correlação de -1,0 indica que a relação entre X e Y é linear perfeitamente negativo.

$$r_{xy} = \frac{\frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n-1}}{\left(\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \right) \cdot \left(\sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}} \right)} \quad (10)$$

Em outras palavras, uma correlação de 1,0 indica que quando a variável X aumenta, a variável y aumenta na mesma proporção, ou no mesmo sentido. Quando a correlação é -1,0 e as variáveis X aumenta, a variável Y irá DE-vincos na mesma proporção. Vamos dar os dados da Tabela 1, utilizando um design mistura com 0,2 de resolução para associar a portfólio e correlação. O risco de os portfólios possíveis para essa configuração é mostrado na Tabela 5 e Fig. 1.

Os ativos que irão compor cada cluster no design mistura cruzou aqui apresentados serão aqueles que a correlação é maior negativo.

Tabela 5 - Efeito da correlação no risco

Correlação		-1.000	-0.750	-0.500	-0.250	0.000	0.250	0.500	0.750	1.000
w_1	w_2	Risco								
0.000	1.000	0.227	0.227	0.227	0.227	0.227	0.227	0.227	0.227	0.227
0.200	0.800	0.143	0.155	0.166	0.176	0.185	0.195	0.203	0.212	0.220
0.400	0.600	0.059	0.094	0.118	0.138	0.156	0.172	0.187	0.200	0.213
0.600	0.400	0.024	0.076	0.105	0.127	0.146	0.163	0.179	0.193	0.206
0.800	0.200	0.108	0.123	0.136	0.149	0.160	0.170	0.180	0.190	0.199
1.000	0.000	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192	0.192

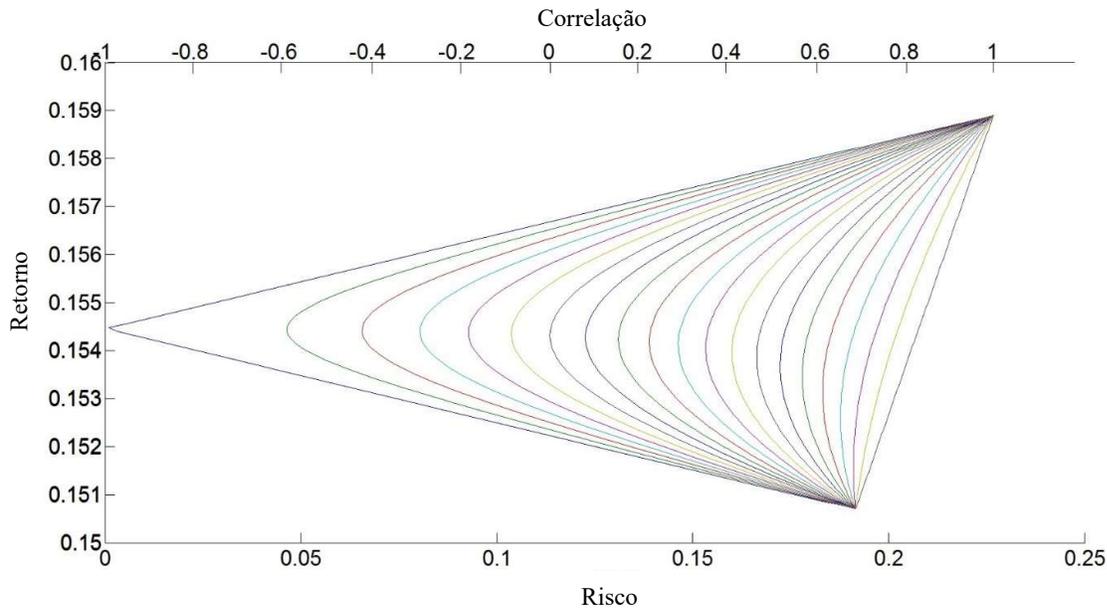


Figura 1 - Efeito da correlação no risco

3.2. Aplicação numérica

O MDE suporta limitações individuais e globais. Considerando-se a resolução da mistura como V , η qualquer número natural inteiro, Ψ o máximo valor da mistura e n o número de ativos por *cluster* ou portfólio da mistura global. Quando V diminui, o número de execuções no experimento aumenta (na Tabela 1 é possível ver isso), $V = 0.5$ e $n = 3$ (11).

$$V = \eta_i \cdot \Pi_{\eta_1} \cdot V = \Psi \quad e \quad \Pi_{\eta_2} \cdot V = n \quad (11)$$

Veja em Cornell (2011b), cada componente pode ser constringido no nível inferior, superior ou em ambos os lados generalizados em (13). Isso significa que o investidor, pode escolher o intervalo da proporção de investir em um ou mais ativos, ou portfólios.

$$\begin{aligned} 0.55 &\leq c_{11} \leq 0.80 \\ 0.30 &\leq c_{12} \leq 0.50 \\ 0.10 &\leq c_{13} \leq 0.40 \end{aligned} \quad (12)$$

$$0 \leq L_i \leq x_i \leq U_i \leq 1.0, \quad 1 < i < q \quad (13)$$

Os dados foram coletados nos sites oficiais do Nasdaq-EUA e BM&FBovespa-BRA (Brasil). Ele inclui 115 meses entre Junho/2004 e Dezembro/2013, de valores de fechamento de 24 empresas aleatórios escolhidos (Tabela 6).

Tabela 6 - Códigos e nomes das empresas selecionadas

Nasdaq-EUA		BM&FBovespa-BRA	
Código	Nome	Código	Nome
AA	Alcoa	BBAS3	Bco Brasil
AAPL	Apple	BBDC4	Bco Bradesco
MSFT	Microsoft	AMBV3	AMBEV
F	Ford	GGBR4	Gerdau
HA	Hawaiian	POMO4	Marcopolo
K	Kellogg	ITUB4	Bco Itau
L	Loews	ALPA4	Alpargatas
M	Macy's	PETR3	Petrobras
RAD	Rite	PETR4	Petrobras
SA	Seabridge	USIM5	Usiminas
TAC	TransAlta	VALE3	Vale
UAM	Universal	VALE5	Vale

Os componentes das séries de dados são considerados determinista, porque de acordo Granger & Newbold (1977), em uma sequência temporal M_t na qual exista uma função de passado e futuro $g_t = g(M_{t-j}, j = 0, 1, \dots)$ tal que (14), significa que os dados são determinísticos.

$$E[(M_{t+j} - g_t)^2] = 0 \quad (14)$$

Ele também demonstra que, se a função g_t é uma função linear de $M_{t-j}, j \geq 0$, M_t será um fator determinístico linear. Usando dados determinísticos em experimentos simulados não vai leva o resultado a uma possível medida de um erro. Para resolver o problema de os dados determinísticos, foi escolhido o método de *bootstrap* visto em Efron (1979). *Bootstrap* é um método amplamente aplicável que torna possível ter algumas estimativas que seriam difíceis para obter através de outros métodos (Wehrens *et al.*, 2000). No essencial, é um conjunto de dados aleatoriamente re-amostrados usando substituição (Henderson, 2005). A estimativa do erro padrão para a média \bar{x} baseada em n pontos independentes dos dados x_1, x_2, \dots, x_n é dado por (15), na qual S^2 é a variância da amostra.

$$SSE = \sqrt{\frac{S^2}{n}} \quad (15)$$

Efron & Tibshirani (1994) mostra que o erro padrão do *bootstrap* é o desvio padrão da replicação. Em (16), onde $s(\cdot) = \sum_{b=1}^B s(x^{*b})/B$ com $s(x) = \bar{x}$.

$$\hat{s}_{e_{boot}} = \left\{ \sum_{b=1}^B [s(x^{*b}) - s(\cdot)]^2 / (B - 1) \right\}^{\frac{1}{2}} \quad (16)$$

Para calcular o retorno μ e o risco σ^2 de cada ativo, foi aplicada a função *log* (17), onde P_t é o mês atual P_{t-1} é o mês anterior.

$$r_i = \frac{\log(P_i)}{\log(P_{i-1})} \quad (17)$$

A equação problema a CM-DOE aplicada na otimização de portfólio pode ser escrita como uma função multiobjectivo em que o retorno deve ser maximizado e o risco deve ser minimizado (18).

$$\begin{aligned} \underset{E, \sigma_p^2}{\text{minimize}} \quad & -f(E) = \sum_{i=1}^n w_i \cdot \gamma_i^T \\ f(\sigma_p^2) = & \sum_{i=1}^n w_i^2 (\gamma_i^T)^2 \sigma_{w\gamma} \end{aligned} \tag{18}$$

sujeito a:

$$\sum_{i=1}^n w_i x_i = \rho$$

RESULTADOS E DISCUSSÕES

Para demonstrar a proposta, vamos considerar três portfólios menores ou *clusters* com três ativos de cada um. O número de ativos por *cluster* não necessita ser o mesmo, estes podem ser diferentes em diferentes grupos. Os valores de γ_n na Tabela 7 são os montantes a investir em cada empresa de cada *cluster*.

Tabela 7 - Pesos individuais nos *clusters*

Cluster 1	γ_1	Cluster 2	γ_2	Cluster 3	γ_3
SA	-0.457	F	0.399	AAPL	-0.993
PETR4	0.439	TAC	-1.000	BBAS3	1.582
VALE5	1.019	USIM5	1.601	ALPA4	0.410

A Fig. 2 representa as fronteiras eficientes de um portfólio mistura cruzada de três clusters. A Fig. 3 mostra as fronteiras eficientes da proposta (CM-DOE) versus a fronteira eficiente usando MDE. Acontece possível ver que os dois modelos são semelhantes. A comparação final pode ser feita usando a Tabela 8, onde os resultados foram considerando quarenta simulação. Os valores mais próximos de 5% de retorno e, em seguida, seu risco relativo das simulações foram extraídos.

Tabela 8 - Finals return and risk

	CM-DOE	MDE
Retorno	0.051	0.052
Risco	0.034	0.047

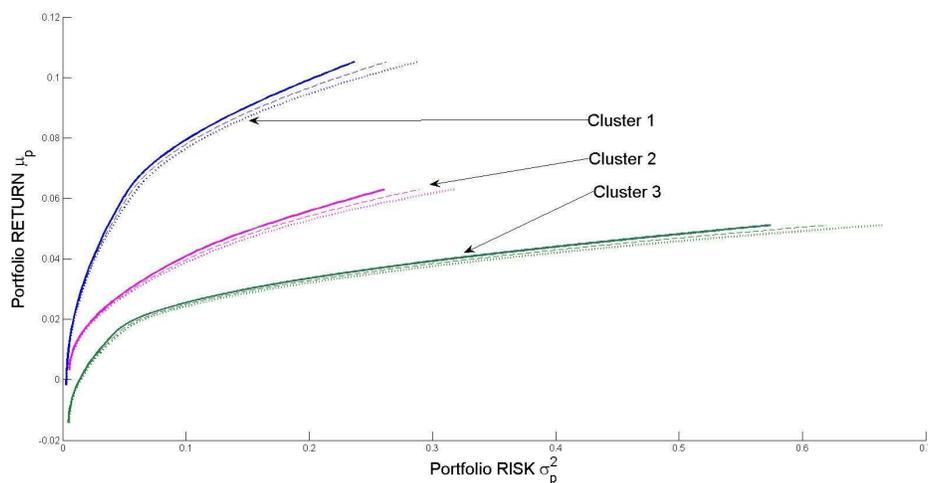


Figura 2 - Fronteira eficientes multi-*cluster*

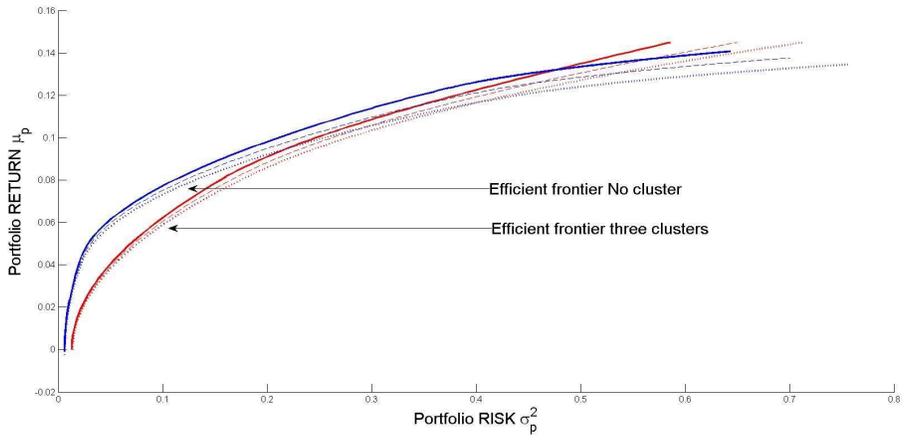


Figura 3 - Fronteira eficiente *clusters* (CM-DOE) x MDE

A correlação entre os grupos não foi considerada neste caso. Em seguida, fizemos outra comparação usando todas as vinte e quatro ativos em um portfólio composto por quatro grupos de seis ativos cada. A consideração final será o impacto da correlação quando se utiliza CM-DOE. Veja a Fig. 4 e 5. Como parece no caso anterior, a mesma ocorrer neste caso. As fronteiras eficientes são semelhantes, não considerando a correlação entre os clusters. O *cluster* pode ser usado também para escolher, em cada um dos ativos, aqueles que a correlação é fortemente negativa e, para o portfólio global, os *clusters* em que a correlação é também fortemente negativa.

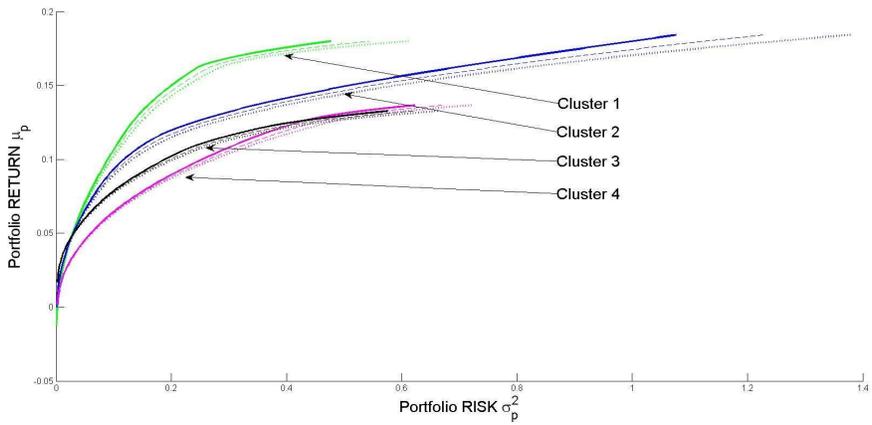


Figura 4 - 24 ativos em 4 *cluster* de 6 ativos cada

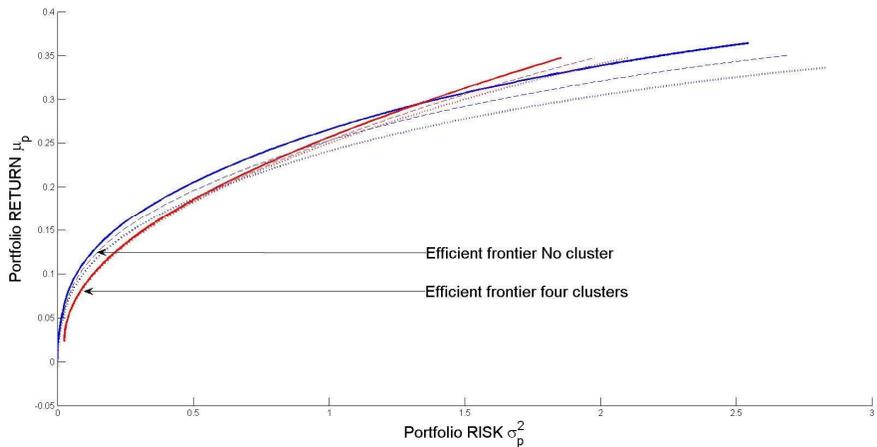


Figura 5 - 24 ativos MDE x 4 *clusters* CM-DOE

CONCLUSÕES

O projeto de experimentos de mistura cruzado (CM-DOE) pode ser aplicado a seleção de portfólios e na sua otimização. Isso leva a um resultado semelhante do método dos polinômios canônicos (MDE) apresentado por De Oliveira *et al.*(2011) e tem menos esforço computacional porque cada mistura terá menos execuções. Em um caso de $n = 3$ e $k = 24$, o total de possibilidades de execuções seria de $3^{23} = 94,143,178,827$. Diferentemente do caso de quatro clusters com $n = 3$ e $k = 6$ cada, o total de execuções seria de apenas $3^5 = 243 \times 4 = 972$ experimentos. Essa foi a principal melhoria da proposta.

A correlação em um *cluster* e entre os *clusters* são equivalentes à correlação entre os ativos escolhidos para compor os *clusters*. A melhor maneira de fazer esta escolha é usando o conhecimento do investidor sobre o mercado para assim selecionar os ativos para compor cada grupo e depois utilizar a correlação para escolher os *clusters* que irão compor o portfólio global.

Para estudos futuros, uma aplicação do CM-DOE na análise multi-período de portfólio e aplicação em produtos reais fora do mercado da Bolsa de Valores deve ser uma ferramenta importante para ajudar as empresas a economizar dinheiro. Nesses casos, o produto que tem muitos componentes e alguns dos seus componentes têm um preço elevado e o número muito grande de execuções será um fator decisivo para a escolha do uso de experimentos.

Agradecimentos

Este trabalho foi apoiado pelas agências brasileiras (CNPq, CAPES, projeto 12875-13-9).

REFERÊNCIAS

- Amihud, Yakov, & Barnea, Amir. 1974. Portfolio selection for managerial control. *Omega*, 2(6), 775 783.
- Andersson, Oivind. 2012. *Experiment!: Planning, Implementing and Interpreting*. John Wiley & Sons.
- Artzner, Philippe, Delbaen, Freddy, Eber, Jean-Marc, & Heath, David. 1999. Coherent measures of risk. *Mathematical Finance*, 9(3), 203 228.
- Bertland, Daniel, & Asgharian, Hossein. 2008. *The Perfect Portfolio*.
- Better, Marco, & Glover, Fred. 2006. Selecting project portfolios by optimizing simulations. *The Engineering Economist*, 51(2), 81 97.
- Borges, Cleber N, Bruns, Roy E, Almeida, Aline A, & Scarminio, Ieda S. 2007. Mixture mixture design for the fingerprint optimization of chromatographic mobile phases and extraction solutions for *Camellia sinensis*. *Analytica chimica acta*, 595(1), 28 37.
- Brandvik, Per Johan, & Daling, Per S. 1998. Optimisation of oil spill dispersant composition by mixture design and response surface methods. *Chemometrics and intelligent laboratory systems*, 42(1), 63 72.
- Cass, David, & Stiglitz, Joseph E. 1970. The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds. *Journal of Economic Theory*, 2(2), 122 160.
- Chan, Ling-Yau. 1988. Optimal design for a linear log contrast model for experiments with mixtures. *Journal of statistical planning and inference*, 20(1), 105 113.
- Chiarawongse, Anant, Kiatsupaibul, Seksan, Tirapat, Sunti, & Roy, Benjamin Van. 2012. Portfolio selection with qualitative input. *Journal of Banking & Finance*, 36(2), 489 496.
- Cornell, John. 2002. *Experiments with mixtures*. 28. New York: John Wiley & Sons, Inc.
- Cornell, John A. 1971. Process variables in the mixture problem for categorized components. *Journal of the American Statistical Association*, 66(333), 42 48.
- Cornell, John A. 2011a. *Experiments with mixtures: designs, models, and the analysis of mixture data*. Vol. 895. John Wiley & Sons.
- Cornell, John A. 2011b. *Experiments with mixtures: designs, models, and the analysis of mixture data*. Vol. 895. John Wiley & Sons.
- Cornell, John A, & Ramsey, Phillip J. 1997. Modeling the component linear and nonlinear blending properties in a two-stage mixture experiment. *Nonlinear Analysis: Theory, Methods & Applications*, 30(7), 4041 4050.

- Costa, OLV, & Paiva, AC. 2002. Robust portfolio selection using linear-matrix inequalities. *Journal of Economic Dynamics and Control*, 26(6), 889 909.
- De Oliveira, Francisco Alexandre, De Paiva, Anderson Paulo, Lima, José Wanderley Marangon, Balestrassi, Pedro Paulo, & Mendes, Ronã Rinston Amaury. 2011. Portfolio optimization using Mixture Design of Experiments: Scheduling trades within electricity markets. *Energy Economics*, 33(1), 24 32.
- Delarue, Erik, De Jonghe, Cedric, Belmans, Ronnie, & D'haeseleer, William. 2011. Applying portfolio theory to the electricity sector: Energy versus power. *Energy Economics*, 33(1), 12 23.
- Di Zio, Marco, Guarnera, Ugo, & Rocci, Roberto. 2007. A mixture of mixture models for a classification problem: The unity measure error. *Computational statistics & data analysis*, 51(5), 2573 2585.
- Draper, Norman R, & Pukelsheim, Friedrich. 1999. Kiefer ordering of simplex designs for first-and second-degree mixture models. *Journal of statistical planning and inference*, 79(2), 325 348.
- Duineveld, CAA, Smilde, AK, & Doornbos, DA. 1993. Designs for mixture and process variables applied in Tabelat formulations. *Analytica chimica acta*, 277(2), 455 465.
- Efron, Bradley. 1979. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 1 26.
- Efron, Bradley, & Tibshirani, Robert J. 1994. *An introduction to the bootstrap*. Vol. 57. CRC press.
- Eriksson, Lennart, Johansson, Erik, & Wikstr m, Conny. 1998. Mixture design design generation, PLS analysis, and model usage. *Chemometrics and Intelligent Laboratory Systems*, 43(1), 1 24.
- Gozálvez-Zafrilla, JM, Santafé-Moros, A, & García-Díaz, J Carlos. 2013. Crossed mixture process design approach to model nano filtration rejection for non-dilute multi-ionic solutions in a given range of solution compositions. *Desalination*, 315, 61 69.
- Granger, Clive William John, & Newbold, Paul. 1977. *Forecasting economic time series*. Economic Theory and Mathematical Economics (USA).
- Henderson, A. Ralph. 2005. The bootstrap: a technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clinica Chimica Acta*, 359(1), 1 26.
- Kang, Lulu, Roshan Joseph, V, & Brenneman, William A. 2011. Design and modeling strategies for mixture-of-mixtures experiments. *Technometrics*, 53(2), 125 136.
- Kennedy, John J, & Bush, Andrew J. 1985. *An introduction to the design and analysis of experiments in behavioral research*. University Press of America.
- Li, Duan, & Ng, Wan-Lung. 2000. Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation. *Mathematical Finance*, 10(3), 387 406.
- Lonni, Audrey Alesandra Stinghen Garcia, Longhini, Renata, Lopes, Gisely Cristiny, de Mello, João Carlos Palazzo, & Scarminio, Ieda Spacino. 2012. Statistical mixture design selective extraction of compounds with antioxidant activity and total polyphenol content from *Trichilia catigua*. *Analytica chimica acta*, 719, 57 60.
- Mage, Ingrid, & Naes, Tormod. 2005. Split-plot design for mixture experiments with process variables: A comparison of design strategies. *Chemometrics and intelligent laboratory systems*, 78(1), 81 95.
- Markowitz, Harry. 1952. Portfolio selection*. *The journal of finance*, 7(1), 77 91.
- Markowitz, Harry. 1959. *Portfolio selection: efficient diversification of investments*. New York.
- Mayers, David. 1974. Portfolio theory, job choice and the equilibrium structure of expected wages. *Journal of Financial Economics*, 1(1), 23 42.
- Mazaheri Tehrani, Mostafa, Saadatmand Tarzjan, Mahdi, *et al*. 2012. Optimization of low-cholesterol-low-fat mayonnaise formulation: Effect of using soy milk and some stabilizer by a mixture design approach. *Food Hydrocolloids*, 28.
- Newbold, Paul, Carlson, William, & Thorne, Betty. 2012. *Statistics for Business and Economics*. Pearson Higher Ed.
- Ngun, Bun Kim, Mohamad, Hasmaliza, Katsumata, Ken-ichi, Okada, Kiyoshi, & Ahmad, Zainal Ari n. 2014. Using design of mixture experiments to optimize triaxial ceramic tile compositions incorporating Cambodian clays. *Applied Clay Science*, 87, 97 107.

- Piepel, Greg F. 1999. Modeling methods for mixture-of-mixtures experiments applied to a Tabelat formulation problem. *Pharmaceutical development and technology*, 4(4), 593 606.
- Scheffé Henry. 1958. Experiments with mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 344 360.
- Scheffé, Henry. 1963. The simplex-centroid design for experiments with mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 235 263.
- Shalit, Haim, & Yitzhaki, Shlomo. 1984. Mean-Gini, Portfolio Theory, and the Pricing of Risky Assets. *The Journal of Finance*, 39(5), 1449 1468.
- Trippi, Robert R. 1989. A decision support system for real estate investment portfolio management. *Information & Management*, 16(1), 47 54.
- Vieira, Antonio Fernando de Castro, & Dal Bello, Luiz Henrique Abreu. 2006. Experimentos com mistura para otimiza^o de processos: uma aplicação com respostas não normais. *Pesquisa Operacional*, 26(3), 605 623.
- Wehrens, Ron, Putter, Hein, & Buydens, Lutgarde. 2000. The bootstrap: a tutorial. *Chemometrics and intelligent laboratory systems*, 54(1), 35 52.
- Wolfinger, Russell D, & Lin, Xihong. 1997. Two Taylor-series approximation methods for nonlinear mixed models. *Computational statistics & data analysis*, 25(4), 465 490.
- Yu, Mei, Wang, Shou-Yang, Lai, Kin Keung, & Chao, X. 2005. Multiperiod portfolio selection on a minimax rule. *Dynamics of continuous discrete and impulsive systems series B*, 12(4), 565.
- Zhang, Chongqi, & Wong, Weng Kee. 2013. Optimal designs for mixture models with amount constraints. *Statistics & Probability Letters*, 83(1), 196 202.
- Zhang, Wei-Guo, Liu, Yong-Jun, & Xu, Wei-Jun. 2012. A possibilistic mean-semivariance-entropy model for multi-period portfolio selection with transaction costs. *European Journal of Operational Research*, 222(2), 341 349.