Universidade do Estado do Rio de Janeiro

Centro de Tecnologia e Ciências

Instituto de Matemática e Estatística

João Francisco Neves

Análise dos fundamentos econômicos de modelos de seleção de carteiras de investimentos

Rio de Janeiro

2017

João Francisco Neves

Análise dos fundamentos econômicos de modelos de seleção de carteiras de investimentos

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Ciências Computacionais, ao Instituo de Matemática e Estatística, da Universidade do Estado do Rio de Janeiro.

Orientadora: Prof.ª Dra. Patrícia Nunes da Silva

Coorientador: Prof. Dr. Carlos Frederico Fragoso de Barros

Rio de Janeiro

2017

CATALOGAÇÃO NA FONTE

UERJ / REDE SIRIUS / BIBLIOTECA CTC-A

N518 Neves, João Francisco.
Análise dos fundamentos econômicos de modelos de seleção de carteiras de investimentos /João Francsico Neves. – 2017.
64f. : il.

Orientadora: Patrícia Nunes da Silva.
Coorientador: Carlos Frederico Fragoso de Barros.
Dissertação (Mestrado em Ciências Computacionais) - Universidade do Estado do Rio de Janeiro, Instituto de Matemática e Estatística.

1.Investimentos - Avaliação - Teses. 2. Títulos (Finanças) - Teses. 3. Risco (Economia) - Teses. 4. Modelagem matemática - Teses. I. Silva, Patrícia Nunes da. II. Barros, Carlos Frederico Fragoso de. III. Universidade do Estado do Rio de Janeiro. Instituto de Matemática e Estatística. IV. Título.

CDU 330.322:519.87

Rosalina Barros *CRB/7 - 4204* - Responsável pela elaboração da ficha catalográfica.

João Francisco Neves

Análise dos fundamentos matemáticos de modelos de seleção de carteiras de investimentos

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação em Ciências Computacionais, da Universidade do Estado do Rio de Janeiro.

Aprovada em 15 de março de 2017.

Banca Examinadora:

____

Prof.ª Dra. Patrícia Nunes da Silva (Orientadora)

Instituto de Matemática e Estatística - UERJ

____

Prof. Dr. Carlos Frederico Fragoso de Barros (Coorientador)

Instituto de Matemática e Estatística - UERJ

____

Prof. Dr. André Luiz Cordeiro dos Santos

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca

____

Prof.ª Dra. Cristiane Oliveira de Faria

Instituto de Matemática e Estatística - UERJ

Rio de Janeiro

2017

DEDICATÓRIA

Dedico esse trabalho a Deus, a minha mãe e ao meu irmão que sempre estiveram ao meu lado me apoiando nos momentos de dificuldades.

AGRADECIMENTOS

Agradeço a Prof.ª Drª Patrícia Nunes da Silva, que sempre esteve disposta a me ajudar nesse trabalho e a todos os meus colegas de turma do mestrado que sempre foram prestativos em me ajudar.

RESUMO

NEVES, João Francisco. *Análise dos fundamentos mamtemáticos de modelos de seleção de carteiras de investimentos.* 2017. xxf. Dissertação (Mestrado em Ciências Computacionais) – Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2017.

A contribuição desse trabalho foi detalhar os fundamentos econômicos relativos a modelos de seleção de carteira de investimentos por meio da modelagem matemática. Os resultados obtidos foram correlacionados com os modelos de otimização de um portfólio baseados nos dois primeiros momentos centrais e nos três primeiros momentos centrais, respectivamente. Nos modelos de seleção de carteira, os portfólios são caracterizados como solução de problemas de otimização condicionados. Estabelecemos uma relação entre o princípio de maximização da função utilidade esperada e os problemas de otimização de um portfólio, a partir de uma análise da função utilidade e de seus parâmetros. Fizemos uma análise da influência da diversificação na otimização do portfólio, de modo que o investidor não deve se ater apenas ao risco presente no ativo, mas, sim, à variação da variância dentro do portfólio causada por esse ativo. Foi demonstrado o benefício obtido pelo investidor ao ter sua riqueza fracionada em diversos ativos financeiros, além da discussão sobre os modelos de otimização baseados em momentos centrais de alta ordem, estabelecendo as vantagens e desvantagens do uso destes. Colocou-se criteriosamente a viabilidade da aplicação dos modelos de otimização em uma carteira de investimento.

Palavras-chave: Utilidade esperada. Otimização de portfólio. Momentos centrais de ordem par e ímpar. Diversificação.

ABSTRACT

NEVES, João Francisco. *Analysis of the economic fundamentals of portfolio selection models.* 2017. f. Dissertação (Mestrado em Ciências Computacionais) – Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2017.

The contribution of this work was to detail the economic fundamentals related to investment portfolio selection models through mathematical modeling. The results were correlated with the optimization models of a portfolio based on the first two central moments and the first three central moments, respectively. In portfolio selection models, portfolios are characterized as solving conditioned optimization problems. We have established a relation between the maximization principle of the expected utility function and the optimization problems of a portfolio, from an analysis of the utility function of its parameters. We have done an analysis of the influence of diversification on portfolio optimization, so that the investor should not only focus on the risk present in the asset, but rather on the variance within the portfolio of the asset. The benefit obtained by the investor was demonstrated by having its fractional wealth in several financial assets, as well as the discussion about the optimization models based on high order central moments, establishing the advantages and disadvantages of their use. The feasibility of applying the optimization models in an investment portfolio was carefully established.

Keywords: Expected utility. Optimization of portfolio. Odd and even central moments. Diversification.

LISTA DE FIGURAS

SUMÁRIO

INTRODUÇÃO

O termo *ativo financeiro* é bastante utilizado na área financeira para designar bens, valores, créditos e direitos que já foram patrimônio de uma determinada pessoa (singular ou coletiva) e que têm na base de sua avaliação os seus respectivos custos. De forma simplificada, os ativos financeiros são aqueles que apenas são negociados no mercado financeiro, sendo conhecidos como "ativos de papel". Existem diversos tipos de ativos e, dentre eles, os mais comuns são as ações, os contratos futuros, os títulos públicos e privados, as commodities, o contrato a termo, as opções e moedas/câmbio.

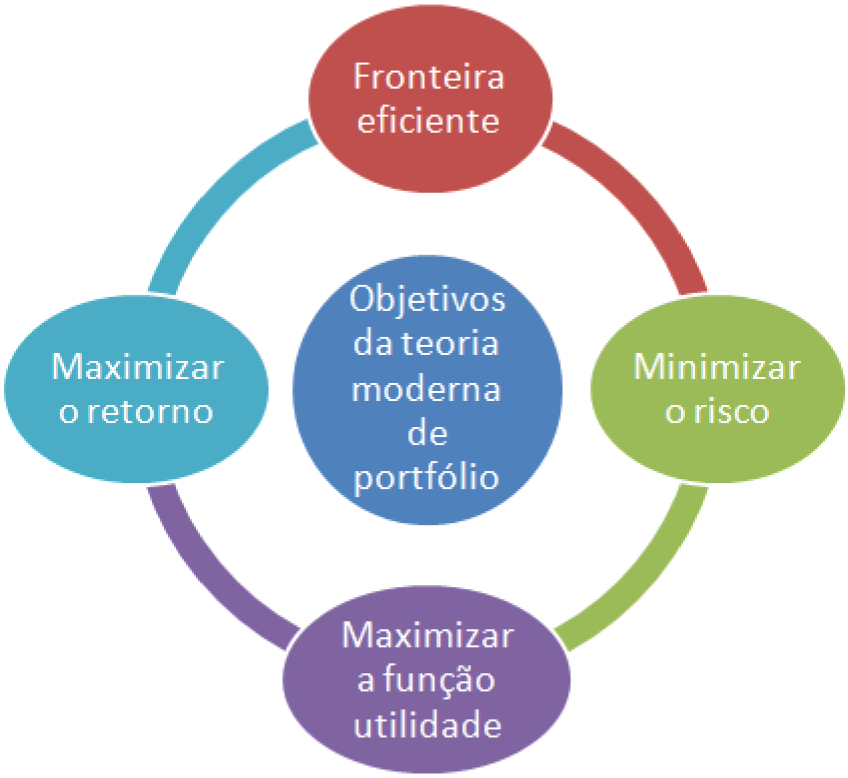
Um portfólio ou carteira de investimento é um conjunto de ativos financeiros que não necessariamente possuem as mesmas características, podendo ser de diversos tipos e de setores econômicos distintos. A análise de portfólio está dividida em duas partes: na análise dos ativos financeiros e na combinação deles dentro de uma carteira de investimentos. A teoria moderna de portfólio lida com a segunda parte e tem por objetivo determinar a fronteira eficiente, isto é, o conjunto de carteiras que minimizam o risco, fixado um nível de retorno ou que maximizam o retorno, fixado um certo nível de risco. Desse modo, é possível, de acordo com as preferências de risco do investidor, determinar a composição ótima de seu portfólio.

Na teoria moderna de portfólio presume-se que o investidor procura maximizar sua função utilidade – na Figura 1 estão ilustrados os objetivos da teoria moderna de portfólio. A fronteira eficiente é composta por um conjunto de portfólios: o investidor escolhe um portfólio que compõe a fronteira eficiente de acordo com o risco e com o retorno de seu interesse, sempre com o intuito de maximizar a função utilidade.

Assume-se, também, que todo investidor apresenta aversão usual ao risco, de modo que a escolha ótima do portfólio depende do risco que o investidor está disposto a correr. Além disso, entende-se que o mercado é perfeito, ou seja, os custos e impostos nas transações financeiras não existem e os ativos financeiros são indefinidamente divisíveis.

Ainda que Markowitz não tenha sido o único nem o primeiro a destacar a importância da diversificação na composição dos portfólios, ele foi pioneiro em modelar matematicamente o princípio de diversificação em investimentos. Rubinstein (2002) afirma que a maior contribuição de Markowitz foi indicar ao investidor que ele não deve se ater ao risco de cada ativo e, sim, à contribuição dele para a variância de todo o portfólio.

Figura 1 – Objetivos da Teoria Moderna de Portfólio.



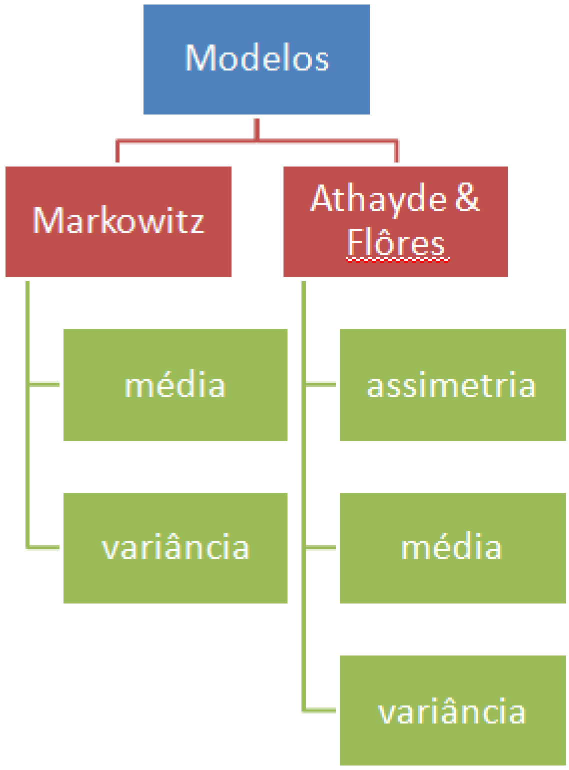
Com a diversificação dos ativos financeiros consegue-se reduzir o risco presente no portfólio mantendo-se a taxa de retorno. Isso ocorre devido à correlação entre os ativos financeiros, que influenciam a variância do portfólio (RUBINSTEIN, 2002).

Samuelson (1967) fez uma análise matemática sobre a importância da diversificação em uma carteira de investimento e demonstrou que é mais vantajoso para o investidor ter a sua riqueza fracionada em diversos ativos financeiros evitando, assim, grandes perdas financeiras. No modelo estabelecido por Markowitz (1952), está necessariamente subjacente uma função utilidade definida a partir de média e de variância. Segundo Samuelson (1970), para que essa função seja uma boa aproximação da função utilidade do investidor, é preciso que sejam considerados alguns fatores. O portfólio ótimo determinado pelo modelo média-variância só corresponde à maximização da função utilidade se a distribuição dos ativos é normal ou se a função utilidade do investidor é quadrática. Samuelson (1970) mostrou que há uma família de parâmetro de distribuições de probabilidade de ativos para a qual o portfólio ótimo do modelo média-variância fornece uma boa aproximação do problema de maximização da função utilidade.

Dados empíricos revelam que média e variância são insuficientes para descrever o comportamento dos ativos, pois são poucos os mercados financeiros que apresentam uma distribuição normal no conjunto de ativos financeiros e, consequentemente, a representação da função utilidade é comprometida (NIELSEN, 2008). Deste dado surge a necessidade de considerar momentos de ordem superior para representar a função utilidade e a fronteira eficiente de forma mais precisa. Dependendo do comportamento da economia, considerando as características de cada mercado financeiro, o modelo média-variância não é o ideal, podendo vir a fornecer um resultado para o problema de otimização pouco preciso. Em tais situações é conveniente utilizar momentos de alta ordem, conseguindo, assim, aumentar a precisão da solução aproximada para o problema de otimização. Isso será discutido elaboradamente mais adiante.

A contribuição deste trabalho será detalhar alguns dos fundamentos econômicos subjacentes a modelos de seleção de carteiras de investimentos através da modelagem matemática. Em uma primeira etapa serão apresentados conceitos teóricos matemáticos necessários à fundamentação da teoria de portfólio e, posteriormente, os modelos propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004). Na Figura 2 ilustra-se que os modelo propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004) são estruturados, respectivamente, pelos dois primeiros momentos centrais e pelos três primeiros momentos centrais. Em seguida, haverá uma exposição detalhada sobre os momentos de ordem par e ímpar presentes na função utilidade. Em seguida, haverá uma exposição detalhada sobre os momentos de ordem par e ímpar presentes na função utilidade. Além disso, será deduzida a preferência do investidor pelos momentos de ordem ímpar e a aversão aos momentos de ordem par (SCOTT & HORVATH, 1980). Por meio da análise dos momentos de ordem par e ímpar será mostrada a motivação para os problemas de otimização propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004). Serão detalhados os resultados de Samuelson (1967) sobre a importância da diversificação em uma carteira de investimento e construída, de forma heurística, uma crítica sobre as aproximações da função utilidade e sobre o uso de modelos baseados apenas nos dois primeiros ou nos três primeiros momentos, como em Markowitz (1952) e Athayde & Flôres Jr. (2004), respectivamente.

Figura 2 – Modelos de otimização de portfólio propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004).



1 **CONCEITOS PRELIMINARES**

Para compreender modelos de seleção de carteiras de investimentos, são necessários os conceitos de função utilidade, valor esperado dessa função e fronteira eficiente, os quais são fundamentados através das teorias de estatística e probabilidade – a fronteira eficiente é obtida pela resolução de um problema de otimização via multiplicadores de Lagrange. Os conceitos e resultados apresentados neste capítulo foram embasados nas referências Bortolossi (2002), Maibaum (1988), Spiegel (1976), Costa Neto (1977), Ash (2000) e Fernandez (2002).

1.1 **Conceitos básicos sobre probabilidade**

Considerar-se-á aqui um espaço de probabilidade $(\Omega, \mathfrak{F}, P)$ com medida um.

1.1.1 Variáveis aleatórias

Definição 1.1.1. *Uma variável aleatória X no espaço de probabilidade $(\Omega, \mathfrak{F}, P)$ é uma função Borel mensurável*[1] *de Ω em $\mathbb{R}$. Isto é,* $X: (\Omega, \mathfrak{F}) \rightarrow (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. *Se X é uma variável aleatória em $(\Omega, \mathfrak{F}, P)$, a medida de probabilidade P_X induzida por X é dada por:*

$$P_X(B) = P\{\omega \in \Omega, X(\omega) \in B\}, \qquad B \in \mathfrak{B}(\mathbb{R}).$$

Definição 1.1.2. *A função de distribuição de uma variável aleatória X é a função $F = F_X$ de $\mathbb{R}$ em* $[0,1]$ *dada por:*

$$F(x) = P\{\omega, X(\omega) \leq x\}, \qquad x \in \mathbb{R}.$$

[1] A definição de uma função Borel mensurável está no Apêndice A, Definição A.0.7.

Os valores $P_X(B)$, $B \in \mathfrak{B}(\mathbb{R})$ caracterizam a variável aleatória X pois fornecem todas as probabilidades dos eventos envolvendo X.

Definição 1.1.3. *A função de distribuição de uma variável aleatória* X é a função $F = F_X$ *para* $\mathbb{R}$ em [0,1] *definida por* $F(x) = P\{\omega : X(\omega) \leq x\}$, *com* x *real.*

Se $a < b$, temos $F(b) - F(a) = P\{\omega : a < X(\omega) \leq b\} = P_X\,(a, b]$. Isto é, F é uma função de distribuição correspondente à medida de Lebesgue-Stieltjes[2] P_X. Em particular, F é não-decrescente e contínua à direita.

Entende-se que a variável aleatória X é absolutamente contínua se, e somente se, existe f, função mensurável de Borel, real e não negativa, tal que:

$$F(x) = \int_{-\infty}^{x} f(t)dt\,, \qquad x \in \mathbb{R}.$$

Compreende-se que f é a função densidade de X. Como $F(x)$ tende a 1 quando x tende a infinito, temos:

$$\int_{-\infty}^{\infty} f(x)dx = 1\,.$$

Se X é absolutamente contínua com função densidade f, segue que:

$$P_X\,(B) = \int_B f(x)\,dx$$

para cada $B \in \mathfrak{B}(\mathbb{R})$.

Se a medida μ é definida por $\mu(B) = \int_B f(x)\,dx$, $B \in \mathfrak{B}(\mathbb{R})$, então $\mu(a, b] = F(b) - F(a), a < b$. Portanto, μ é a medida de Lebesgue-Stieltjes correspondente a F, consequentemente $\mu = P_X$.

[2] A definição de medida de Lebesgue-Stieltjes está no Apêndice A, Definição A.0.4.

Definição 1.1.4. *Um vetor aleatório n-dimensional no espaço de probabilidade $(\Omega, \mathfrak{F}, P)$ é uma função Borel mensurável de Ω* para $\mathbb{R}^n$.

Um vetor aleatório pode ser considerado como uma n-upla $(X_1, \ldots, X_n)$ de variáveis aleatórias. Portanto, a medida de probabilidade induzida pelo vetor aleatório X é definida por $P_X\,(B) = P\{\omega : X(\omega) \in B\}, B \in \mathfrak{B}(\mathbb{R}^n)$.

A função de distribuição de X é a função $F = F_X$ de $\mathbb{R}^n$ em $[0, 1]$, definida por:

$$F(x) = P_X\,(-\infty, x] = P\backslash\{\omega : X_i(\omega) \leq x_i, i = 1, \ldots, n\};$$

F é também chamada de função de distribuição conjunta de $X_1, \ldots, X_n$; F é não-decrescente e contínua à direita em $\mathbb{R}^n$ e P_X é uma medida de Lebesgue-Stieltjes determinada por F.

O vetor aleatório X é dito absolutamente contínuo se, e somente se, existe uma função Borel mensurável f definida em $\mathbb{R}^n$ não negativa, denominada de função densidade ou densidade de X, tal que:

$$F(x) = \int_{(-\infty, x]} f(t)\, dt\,, \qquad x \in \mathbb{R}^n.$$

Segue que $P_X\,(B) = \int_B f(x)\,dx$ para todo $B \in \mathfrak{B}(\mathbb{R}^n)$. Para a medida μ, definida por $\mu\,(B) = \int_B f(x)\,dx\,,\; B \in \mathfrak{B}\,(\mathbb{R}^n)$, temos $\mu\,(a, b] = \int_{(a,b]} f(x)\,dx = F(a, b]$. Segue que μ é a medida de Lebesgue-Stieltjes determinada por F, consequentemente $\mu = P_X$.

1.1.1.1 Variáveis aleatórias independentes

Intuitivamente, dois eventos A e B são independentes se uma declaração concernente à ocorrência ou não ocorrência de um dos eventos não muda a probabilidade do outro evento.

Definição 1.1.5. Sejam $X_1, \ldots, X_n$ variáveis aleatórias no espaço de probabilidade $(\Omega, \mathfrak{F}, P)$, afirma-se que $X_1, \ldots, X_n$ são independentes se, e somente se, para quaisquer $B_1, \ldots, B_n \in \mathfrak{B}(\mathbb{R})$, temos

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\} = P\{X_1 \in B_1\} \cdots P\{X_n \in B_n\}.$$

A independência de $X_1, \ldots, X_n$ pode ser descrita da seguinte maneira:

Teorema 1.1.1. *Sejam $X_1, \ldots, X_n$ variáveis aleatórias em $(\Omega, \mathfrak{F}, P)$. Seja F_i a função de distribuição de X_i, com $i = 1, \ldots, n$, e F a função de distribuição de $X = (X_1, \ldots, X_n)$. Então $X_1, \ldots, X_n$ são independentes se, e somente se:*

$$F(x_1, x_2, \ldots, x_n) = F_1(x_1)F_2(x_2) \cdots F_n(x_n)$$

para todo $x_1, \ldots, x_n \in \mathbb{R}$.

Teorema 1.1.2. *Se $X = (X_1, \ldots, X_n)$ tem densidade f, então cada X_i tem uma densidade f_i. Além disso, nesse caso $X_1, \ldots, X_n$ são independentes se, e somente se, $f(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n)$ para todos $(x_1, \ldots, x_n)$ exceto possivelmente para um subconjunto de Borel com medida de Lebesgue nula.*

1.1.2 Valor esperado

Definição 1.1.6. *Se X é uma variável aleatória em $(\Omega, \mathfrak{F}, P)$, o valor esperado de X é definido por:*

$$E(X) = \int_\Omega X \, dP,$$

desde que a integral exista.

Assim sendo, $E(X)$ é a integral da função Borel mensurável X com relação à medida de probabilidade P. A mesma definição é aplicável se X é uma variável aleatória definida em $\mathfrak{B}(\overline{\mathbb{R}})$.

Teorema 1.1.3. *Sejam X e Y variáveis aleatórias e k uma constante real, serão válidas as seguintes propriedades do valor esperado:*

1) $E(k) = k$;
2) $E(kX) = k \cdot E(X)$;
3) $E(X + Y) = E(X) + E(Y)$;
4) *Se* $X \leq Y$, *então* $E(X) \leq E(Y)$.

Em algumas situações é inconveniente calcular $E(X)$ através da integração sobre Ω. Portanto, será conveniente expressar os resultados de $E(x)$ como uma integral, com respeito à medida de probabilidade induzida P_X, a qual é determinada pela distribuição F.

Teorema 1.1.4. *Sejam X uma variável aleatória em $(\Omega, \mathfrak{F}, P)$ com função de distribuição F e g uma função mensurável de Borel de $\mathbb{R}$ em $\mathbb{R}$. Se $Y = g \circ X$, então:*

$$E(Y) = \int_{\mathbb{R}} g(x)\, dF(x) = \int_{\mathbb{R}} g\, dP_X.$$

As igualdades no Teorema 1.1.4. são entendidas da seguinte maneira: se qualquer uma das integrais está bem definida, então a outra também está e elas coincidem.

Corolário 1.1.1. *(a) Seja X um vetor aleatório em $(\Omega, \mathfrak{F}, P)$ e seja g uma função mensurável de Borel de $\mathbb{R}^n$ em $\mathbb{R}$, então $E(g \circ X) = \int_{\mathbb{R}^n} g(x)\, dF(x)$, desde que a integral exista.*
(b) Se X é uma variável aleatória (ou vetor aleatório) com densidade f, então $\int g(x)\, dF(x) = \int g(x) f(x)\, dx$ (a integração é feita sobre $\mathbb{R}$, se X é uma variável aleatória, ou sobre $\mathbb{R}^n$, se X é um vetor aleatório).

Definição 1.1.7. (Momento). *Seja n um número natural, o n-ésimo momento de uma variável aleatória X é definido como:*

$$\mu_n(X) = E(X^n). \tag{1.1}$$

A média de X corresponde ao momento de ordem um, isto é, para $n = 1$ temos a média:

$$\mu_1 (X) = E(X).$$

Definição 1.1.8. *Seja X uma variável aleatória em $(\Omega, \mathfrak{F}, P)$. Se $k \in \mathbb{N}$, o número $E(X^k)$ é chamado o k-ésimo momento de X; $\mu_k = E\left[\left(X - E(X)\right)^k\right]$ é denominado o k-ésimo momento central. Os momentos centrais são definidos somente quando $E(X)$ é finito.*

A variância ou momento central de ordem dois é definida por:

$$\sigma^2 (X) = \mu_2 = E\left[\left(X - E(X)\right)^2\right]. \tag{1.2}$$

O resultado que segue sobre o valor esperado de um produto de variáveis aleatórias independentes é consequência direta do teorema de Fubini.

Teorema 1.1.5. *Seja $X_1, \dots, X_n$ variáveis aleatórias independentes em $(\Omega, \mathfrak{F}, P)$. Se todos os X_i são não-negativos ou se $E(X_i)$ é finito para todo i, então $E(X_1 \cdots X_n)$ existe e equivale a $E(X_1)E(X_2)\cdots E(X_n)$, isto é:*

$$E(X_1 \cdots X_n) = \int_{\mathbb{R}} x_1 dP_{x_1}(x_1) \cdots \int_{\mathbb{R}} x_n \, dP_{x_n}(x_n) = E(X_1) \cdots E(X_n).$$

Quando duas variáveis X e Y não são independentes, geralmente é interessante avaliar o quanto essas variáveis estão relacionadas. Uma medida do grau e do sinal da correlação linear é dada pela covariância entre as duas variáveis e, por meio da covariância, podemos analisar a dispersão dos valores da variável (X, Y) em relação à $(E(X), E(Y))$.

Definição 1.1.9. *Sejam X e Y variáveis aleatórias com valor esperado finito. Suponha que $E(XY)$ seja finito. A covariância de X e Y é definida por*

$$Cov(X,Y) = E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] = E(XY) - E(X)E(Y).$$

A covariância é positiva ($cov(X,Y) > 0$) se as duas variáveis tendem à mesma direção, isto é, valores de X acima da sua média estão relacionados aos valores de Y acima da sua média, com o mesmo ocorrendo para os valores de ambas as variáveis que estão abaixo das suas respectivas médias. A covariância será negativa ($cov(X,Y) < 0$) se valores acima da média de uma variável estão relacionados aos valores de outra variável que estão abaixo da média. No caso em que a covariância é nula ($cov(X,Y) = 0$), as variáveis X e Y são independentes.

Pode-se estabelecer uma relação entre variância e covariância. Sejam $X_1, X_2, \dots, X_n$ variáveis aleatórias e R a combinação linear dessas variáveis com as constantes $\alpha_1, \alpha_2, \dots, \alpha_n$, isto é, $R = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n$, onde $\alpha_i \in \mathbb{R}$, para $i = 1, \dots, n$.

Pelo Teorema 1.1.3., o valor esperado de R é dado por:

$$E(R) = E(\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n) = \alpha_1 E(X_1) + \alpha_2 E(X_2) + \cdots + \alpha_n E(X_n).$$

A variância de R pode ser obtida a partir da equação (1.2). Isto é, $\sigma^2(R) = E\left[\big(R - E(R)\big)^2\right]$.

Portanto,

$$\sigma^2(R) = E\left[\Big(\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n - \big(\alpha_1 E(X_1) + \alpha_2 E(X_2) + \cdots + \alpha_n E(X_n)\big)\Big)^2\right].$$

Então,

$$\sigma^2(R) = E\left[\Big(\alpha_1\big(X_1 - E(X_1)\big) + \alpha_2\big(X_2 - E(X_2)\big) + \cdots + \alpha_n\big(X_n - E(X_n)\big)\Big)^2\right].$$

Ao se desenvolver os produtos temos:

$$\begin{aligned}\sigma^2(R) = E\Big[\alpha_1^2\big(X_1 - E(X_1)\big)^2 &+ \cdots + \alpha_n^2\big(X_n - E(X_n)\big)^2 \\ &+ 2\alpha_1\alpha_n\big(X_1 - E(X_1)\big)\big(X_n - E(X_n)\big) + \cdots \\ &+ 2\alpha_{n-1}\alpha_n\big(X_{n-1} - E(X_{n-1})\big)\big(X_n - E(X_n)\big)\Big].\end{aligned}$$

.

Como $E\left[\left(X_i - E(X_i)\right)^2\right]$ é o momento central de ordem dois da variável X_i e $E\left[\left(X_i - E(X_i)\right)\left(X_j - E\left(X_j\right)\right)\right]$ é a covariância entre as variáveis X_i e X_j, concluímos que:

$$\sigma^2(R) = \sum_{i=1}^{n} \alpha_i^2\, \sigma^2(X_i) \; + \; 2\sum_{i=1}^{n}\sum_{i>j}^{n} \alpha_i\alpha_j cov(X_i, X_j).$$

Quando $X_1, \ldots, X_n$ são independentes, $Cov(X_i, X_j) \; = \; 0$ para $i \neq \; j$, portanto:

$$\sigma^2(R) \; = \; \sum_{i=1}^{n} \alpha_i^2 \sigma^2(X_i).$$

Não obstante, calcular a variância de X_i é equivalente a calcular a covariância de X_i e X_i, isto é, $\sigma^2(X_i) \; = \; E\left[\left(X_i - E(X_i)\right)^2\right] \; = \; E\left[\left(X_i - E(X_i)\right)\left(X_i - E(X_i)\right)\right] \; = \; cov(X_i, X_i)$. Logo, a variância da variável aleatória R pode ser calculada por:

$$\sigma^2(R) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j \; cov\left(X_i, X_j\right). \tag{1.3}$$

1.2 **Otimização com restrições**

O método de otimização exposto será para o caso em que o conjunto admissível é construído com o uso de várias restrições na forma de equações (BORTOLOSSI, 2002). Mais precisamente, será analisado como encontrar os extremos de uma função f de n variáveis no conjunto admissível

$$D \; = \; \{x \in \; \mathbb{R}^n | h_1(x) = c_1, \ldots, h_m(x) = c_m\}$$

formado por todos os pontos x de $\mathbb{R}^n$ que satisfazem, simultaneamente, as m restrições $h_1(x) = c_1, \dots, h_m(x) = c_m$. Note que o conjunto D nada mais é que o conjunto de nível da função vetorial $h(x) = (h_1(x), \dots, h_m(x))$ de $\mathbb{R}^n$ em $\mathbb{R}^m$ associado ao nível $c = (c_1, \dots, c_m)$:

$$D = \mathcal{F}_c = \{\, x \in \mathbb{R}^n \mid h(x) \;=\; \big(h_1(x), \dots, h_m(x)\big) = (c_1, \dots, c_m) = c\}.$$

Teorema 1.2.1. (Multiplicadores de Lagrange) *Sejam $f, h_1, \dots, h_m$ funções de classe C^1 de n variáveis e seja p um extremo (máximo ou mínimo) local de f no conjunto admissível*

$$D = \{x \in \mathbb{R}^n \mid h_1(x) = c_1, \dots, h_m\,(x) = c_m\}.$$

Suponha que p satisfaça a seguinte condição de regularidade: o posto da matriz Jacobiana

$$\begin{bmatrix} \nabla h_1(p) \\ \vdots \\ \nabla h_m(p) \end{bmatrix}_{m\,x\,n} = \begin{bmatrix} \dfrac{\partial h_1}{\partial x_1}(p) & \cdots & \dfrac{\partial h_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_m}{\partial x_1}(p) & \cdots & \dfrac{\partial h_m}{\partial x_n}(p) \end{bmatrix}_{m\,x\,n}$$

é igual a m (o número de restrições). Então, existem números reais $\lambda_1^*, \dots, \lambda_m^*$ (os multiplicadores de Lagrange), tais que:

$$\begin{cases} \nabla f(p) = \lambda_1^* \cdot \nabla h_1(p) + \cdots + \lambda_m^* \cdot \nabla h_m(p) \\ h_1(p) = c_1 \\ \vdots \\ h_m(p) = c_m. \end{cases}$$

Isto é, o vetor gradiente $\nabla f(p)$ é uma combinação linear dos vetores gradientes $\nabla\, h_1\,(p), \dots, \nabla\, h_m\,(p)$. O sistema acima é denominado de condição de primeira ordem para o problema de otimização. Equivalentemente, o ponto $(p, \lambda^*) = (p_1, \dots, p_n, \lambda_1^*, \dots, \lambda_m^*)$ é ponto crítico do lagrangeano

$$L(x, \lambda) = f(x) - \lambda_1 \cdot\ [h_1\,(x) -\ c_1]\ - \cdots - \lambda_m \cdot\ [h_m\,(x)\ -\ c_m],$$

isto é,

$$\frac{\partial L}{\partial x_1}(p,\lambda^*) = 0, \dots, \frac{\partial L}{\partial x_n}(p,\lambda^*) = 0,$$
$$\frac{\partial L}{\partial \lambda_1}(p,\lambda^*) = 0, \dots, \frac{\partial L}{\partial \lambda_m}(p,\lambda^*) = 0.$$

A condição de regularidade não pode ser omitida no teorema dos multiplicadores de Lagrange – para omiti-la, é necessário acrescentar um multiplicador para a função-objetivo.

Teorema 1.2.2. (Fritz John) *Sejam $f, h_1, \dots, h_m$ funções de classe C^1 de n variáveis e seja p um extremo (máximo ou mínimo) local de f no conjunto admissível*

$$D = \{x \in \mathbb{R}^n \mid h_1(x) = c_1, \dots, h_m(x) = c_m\}.$$

Então, existem multiplicadores $\lambda_0^*, \dots, \lambda_m^*$, tais que:

$$\begin{cases} \lambda_0^* \cdot \nabla f(p) = \lambda_1^* \cdot h_1(p) + \cdots + \lambda_m^* \cdot h_m(p) \\ h_1(p) = c_1, \\ \vdots \\ h_m(p) = c_m, \\ (\lambda_0^*, \dots, \lambda_m^*) \neq (0, \dots, 0), \\ \lambda_0^* \in \{0, 1\}. \end{cases}$$

A condição $(\lambda_0^*, \dots, \lambda_m^*) \neq (0, \dots, 0)$ diz que os multiplicadores $\lambda_0^*, \dots, \lambda_m^*$ não são simultaneamente nulos.

A hipótese de regularidade do teorema dos multiplicadores de Lagrange garante que seja possível tomar $\lambda_0^* = 1$ no teorema de Fritz John.

2 MODELOS DE SELEÇÃO DE CARTEIRAS DE INVESTIMENTOS

Os modelos de seleção de carteiras de investimentos apresentados nesse capítulo foram propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004). Ambos os modelos têm como objetivo a escolha de um portfólio ótimo dentre o conjunto de portfólios da fronteira eficiente, de acordo com a preferência do investidor sobre o risco e o retorno esperados.

Primeiramente, abordaremos o modelo de Markowitz e, posteriormente, o modelo de Athayde \& Flôres. Serão analisados de forma crítica os dois modelos, buscando contribuir com a teoria moderna de portfólio.

2.1 Modelo de Markowitz

O método proposto por Markowitz de análise de portfólio é construído sobre dois parâmetros: o valor esperado do retorno dos ativos financeiros, isto é, a média do retorno, e o risco presente no portfólio, que é associado à variância/desvio padrão dos ativos financeiros.

Considere uma carteira formada por N ativos financeiros. O objetivo da análise de portfólio é encontrar combinações eficientes desses ativos financeiros, ou seja, determinar qual é a fração α_i do capital que será investida no i-ésimo ativo financeiro, de modo a garantir um compromisso entre o retorno esperado e o risco assumido.

O retorno esperado do portfólio é calculado com a média ponderada dos retornos esperados dos ativos financeiros.

Definição 2.1.1. *O retorno esperado de um portfólio é definido por:*

$$\overline{R_p} = \sum_{i=1}^{N} (\alpha_i \overline{R_i}),$$

onde $\overline{R_i}$ é o retorno esperado do i-ésimo ativo financeiro e α_i é a fração do capital aplicada no i-ésimo ativo financeiro.

A fórmula do retorno esperado de um portfólio pode ser expressa em notação matricial:

$$\overline{R_p} = \alpha\bar{R} = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \cdot \begin{bmatrix} \overline{R_1} \\ \overline{R_2} \\ \vdots \\ \overline{R_n} \end{bmatrix},$$

onde α é o vetor de pesos dos diferentes ativos financeiros e $\bar{R}$ é o vetor do retorno esperado dos ativos financeiros.

Definição 2.1.2. *A variância do portfólio σ_p^2 pode ser obtida a partir da fórmula*

$$\sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j cov_{ij},$$

onde cov_{ij} é a covariância entre o i-ésimo e o j-ésimo ativo financeiro.

Do mesmo modo que o retorno esperado de um portfólio, a variância também pode ser expressa em notação matricial:

$$\sigma_p^2 = \alpha\, M_2 \alpha^t = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \cdot \begin{bmatrix} cov_{11} & \cdots & cov_{1n} \\ \vdots & \ddots & \vdots \\ cov_{n1} & \cdots & cov_{nn} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix},$$

onde M_2 é a matriz de covariância.

Markowitz (1952) afirma que o portfólio ótimo corresponde à solução do problema de minimização da variância σ_p^2, sujeita a duas restrições: um determinado nível de retorno esperado $\overline{R_p^*}$ e a soma dos pesos dos ativos financeiros ser igual a 1, isto é, $\alpha^t\ [1] = 1$, em que [1] é um vetor cujos elementos são todos iguais a um.

A fronteira eficiente é o conjunto de valores ótimos (soluções do problema de otimização) para diferentes valores fixados de retorno, pois o que é pretendido quando se constrói a fronteira eficiente é encontrar qual portfólio, dentro dessa fronteira, tem um mínimo de variância, dado um nível de retorno esperado de acordo com a preferência do investidor.

O investidor tem conhecimento sobre o retorno esperado e o desvio padrão da distribuição dos retornos desses ativos (MARKOWITZ, 1952). Entretanto, como o investidor poderia formar um portfólio a partir desse conjunto de informações que ele possui sobre os ativos?

Markowitz considera que o objetivo do investidor não deve ser apenas o de maximizar o seu retorno esperado dentro de um portfólio; os investidores devem, também, diversificar seus investimentos seguindo, assim, uma regra de ouro da economia. A diversificação não faz sentido se o objetivo é apenas obter o retorno esperado máximo do portfólio. Caso fosse apenas esse o objetivo, os investidores teriam portfólios compostos por um único ativo financeiro: aquele que tem o maior retorno esperado.

Percebe-se, então, a importância da medida de dispersão – a variância –, pois esta medida fornece informação sobre o risco envolvido no conjunto de ativos financeiros dentro de um portfólio. De um modo geral, o investidor prefere o retorno esperado e, em contrapartida, não prima pela variância, por essa medida representar o risco envolvido no investimento. Portanto, a preferência do investidor está sobre um portfólio que maximize o retorno esperado, dada uma determinada variância, ou miniminize a variância dado um determinado retorno esperado.

O critério para estabelecer a análise sobre o portfólio depende da natureza do investidor, pois cada investidor possui diferentes metas a alcançar em seus investimentos. Porém, Markowitz (1952) defende que o modelo proposto por ele pode ser utilizado por todos os tipos de investidores. Isso ocorre devido às características em comum existentes entre os investidores.

O investidor possui uma inclinação a ganhar sempre o máximo possível, isto é, sempre busca maximizar seus retornos. O modelo de Markowitz não é adequado para investidores especuladores, pois ele assume a premissa de que a preferência do investidor seja por investimentos que apresentem retorno estável. Como ativos financeiros com retorno instável são comumente as opções de investimentos dos especuladores, devido à possibilidade de grandes retornos, um investidor com perfil especulador obteria uma análise incoerente com os seus objetivos sobre a combinação de ativos do portfólio por meio do modelo de Markowitz.

Decorre do modelo média-variância de Markowitz que a distribuição dos retornos dos ativos financeiros é normal, ou que a função utilidade do investidor é quadrática.

Uma vez que é assumido que a distribuição dos retornos esperados se comporta como uma distribuição normal, os dois parâmetros média e variância são suficientes para caracterizar tal distribuição. A função utilidade quadrática é uma boa aproximação em relação a outras funções utilidades; tal escolha para essa função não provém de uma observação empírica do comportamento do investidor: ela foi feita por se adequar ao modelo proposto por Markowitz (NIELSEN, 2008).

2.1.1 Crítica sobre a Teoria Moderna de Portfólio

A teoria moderna de portfólio vem sendo amplamente discutida ao longo dos anos. Atualmente, existem diversos modelos de análise de portfólio, como o modelo dinâmico de análise de portfólio, que é mais adequado à realidade do investidor, pois considera mais de um período de investimento. Não obstante, diante de diversos modelos de análise de portfólio, é interessante explanar sobre quando o modelo média/variância é suficiente para selecionar um portfólio. O artigo de Samuelson (1970) pode ser útil para essa discussão.

Se a análise do retorno esperado de um portfólio for feita em cima da média e da variância, apenas, são assumidas pelos menos duas premissas: a distribuição dos retornos dos ativos financeiros é normal, sendo assim, com a média e variância tem-se informação suficiente para descrever toda a distribuição; e a função utilidade quadrática é suficiente para auxiliar nas decisões do investidor, isto é, os momentos de alta ordem são desconsiderados na tomada de decisão.

Muitos estudos mostram que a distribuição dos retornos dos ativos, frequentemente, não é simétrica, apresentando em muitos casos assimetria, principalmente assimetria positiva. E, também, foi observada significativa curtose em mercados financeiros, como o Japonês. Somente uma pequena fração dos mercados financeiros apresentou, na distribuição dos retornos dos ativos financeiros, distribuição normal (NIELSEN, 2008).

Quando o investidor assume a função utilidade quadrática, presume-se que ele não se importa com os momentos de alta ordem. No entanto, na prática, os investidores se importam com a assimetria, momento de terceira ordem, do retorno. Não obstante, o investidor tolera uma leve diminuição da média do retorno esperado, podendo vir a se proteger contra perdas e mantendo a possibilidade de ganhos – isto se refere a uma assimetria positiva –, sendo percebida, portanto, a importância da assimetria na análise de um portfólio.

2.2 **Modelo de Athayde & Flôres Jr.**

Os economistas sabem que uma distribuição dos retornos dos ativos financeiros raramente se comporta como uma distribuição normal, entretanto, ainda assim, permanecem relutantes em incorporar os momentos de alta ordem em suas análises. Geralmente, os investidores preferem valores altos para os momentos ímpares e valores baixos para os momentos de ordem par. Essa é uma forma de fazer decair as perdas extremas no investimento e aumentar o retorno do investimento.

Baseado na premissa de que a direção das preferências dos investidores é positiva (negativa) para valores positivos (negativos) de momentos centrais de ordem ímpar (tal como a assimetria) e negativa para valores positivos de momentos centrais de ordem par (tais como a variância e a curtose), Athayde & Flôres (2004) propuseram um modelo de seleção de carteiras no caso de n ativos de risco e um sem risco, considerando os três primeiros momentos, sendo permitidas vendas a descoberto.

2.2.1 Notação

A assimetria do portfólio pode ser analisada de forma semelhante à variância do portfólio. Entretanto, a assimetria do portfólio não é apenas uma média ponderada do momento de terceira ordem do retorno dos ativos financeiros. A variação dos retornos de um ativo financeiro está sujeita à influência dos outros ativos financeiros, isto é, não se pode analisar os ativos de forma independente, como se não houvesse correlação entre eles. Portanto, para calcular a assimetria do portfólio, não basta calcular apenas a assimetria dos retornos: é necessário calcular a co-assimetria entre os retornos dos ativos financeiros. A co-assimetria é uma medida curvilínea da interação que ocorre dentro de um conjunto estatístico de distribuições do retorno de ativos (NIELSEN, 2008).

Definição 2.2.1 (Co-assimetria entre ativos financeiros)

$$Co-assimetria\left(R_i, R_j, R_k\right) = E\left[(R_i - \overline{R_i})\left(R_j - \overline{R_j}\right)(R_k - \overline{R_k})\right], \qquad (2.1)$$

onde $\overline{R_*}$ é o retorno esperado do ativo $*$.

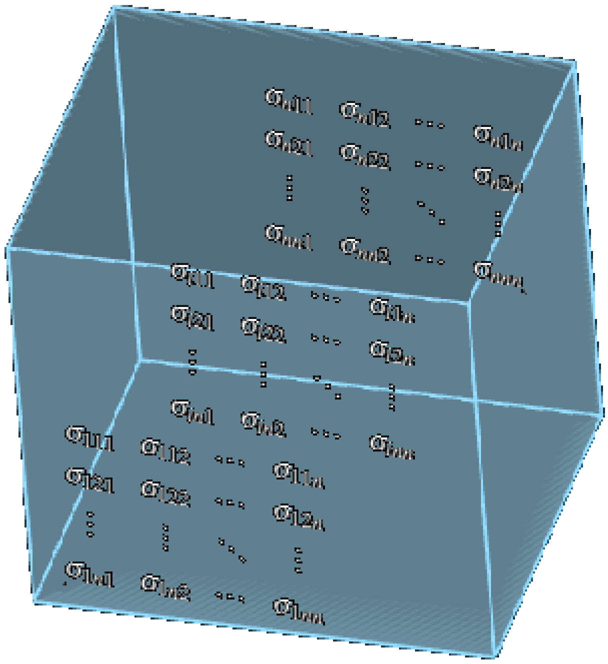
A assimetria do portfólio é calculada do seguinte modo:

$$\sigma_{p^3} = E\left[R_p - \overline{R_p}\right]^3 = \sum_{i,j,k}^{N} \alpha_i\, \alpha_j\, \alpha_k E\left[(R_i - \bar{R}_i)\left(R_j - \bar{R}_j\right)(R_k - \overline{R_k})\right], \qquad (2.2)$$

onde R_p é o retorno atual do portfólio, $\overline{R_p}$ é o retorno esperado do portfólio, α_i, α_j e α_k são respectivamente, os pesos dos ativos i, j e k no portfólio.

Lidar com momentos de alta ordem é algebricamente complicado ou, até mesmo, inviável. Dado um espaço n-dimensional com vetores aleatórios, esses momentos podem ser vistos como um tensor. O tensor do segundo momento é uma matriz $n \times n$ de covariância, enquanto o do terceiro momento pode ser visualizado como uma matriz $n \times n \times n$, um cubo no espaço tridimensional. Na Figura 3 é retratado um exemplo de um tensor do terceiro momento.

Figura 3 – Tensor do momento central de ordem três.



Pode-se transformar o tensor da assimetria em uma matriz $n \times n^2$, dividindo o cubo em n camadas $n \times n$ e colocando-as lado a lado, respeitando a mesma ordem de divisão.

Denotando σ_{ijk} a co-assimetria entre os ativos i, j e k, no caso em que $n = 2$, tem como resultado uma matriz 2×4, como segue:

$$\begin{bmatrix} \sigma_{111} & \sigma_{112} & \sigma_{211} & \sigma_{212} \\ \sigma_{121} & \sigma_{122} & \sigma_{221} & \sigma_{222} \end{bmatrix},$$

onde apenas no máximo quatro elementos são distintos, uma vez que $\sigma_{112} = \sigma_{121} = \sigma_{211}$ e $\sigma_{212} = \sigma_{122} = \sigma_{221}$.

No caso em que há n ativos financeiros, a matriz de co-assimetria é dada por:

$$M_3 = \begin{bmatrix} \sigma_{111} & \sigma_{112} & \cdots & \sigma_{11n} & \cdots & \sigma_{i11} & \sigma_{i12} & \cdots & \sigma_{i1n} & \cdots & \sigma_{n11} & \sigma_{n12} & \cdots & \sigma_{n1n} \\ \sigma_{121} & \sigma_{122} & \cdots & \sigma_{12n} & \cdots & \sigma_{i21} & \sigma_{i22} & \cdots & \sigma_{i2n} & \cdots & \sigma_{n21} & \sigma_{n22} & \cdots & \sigma_{n2n} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n1} & \sigma_{1n2} & \cdots & \sigma_{1nn} & \cdots & \sigma_{in1} & \sigma_{in2} & \cdots & \sigma_{inn} & \cdots & \sigma_{nn1} & \sigma_{nn2} & \cdots & \sigma_{nnn} \end{bmatrix}$$

Considere um vetor de pesos $\alpha \in \mathbb{R}^n$ e dados x, M_2 e M_3 as matrizes que contêm respectivamente o retorno médio excedente[3], covariância e a assimetria associados aos n ativos. O retorno médio excedente, variância e assimetria do portfólio com esses pesos serão dados respectivamente por $\alpha^t x$, $\alpha^t M_2 \alpha$ e $\alpha^t M_3 (\alpha \otimes \alpha)$, onde $\otimes$ é o produto de Kronecker. Note que, α é um vetor com n coordenadas, e as três expressões são funções homogêneas com grau igual à ordem correspondente de cada momento.

2.2.2 Portfólio de variância mínima

O problema de otimização do modelo retorno médio/variância consiste em encontrar o portfólio em que se tem o maior retorno esperado, dada uma determinada variância, ou, de modo correspondente, encontrar a menor variância dado um determinado retorno médio. Já no espaço retorno médio $\times$ variância $\times$ assimetria, dois parâmetros são fixados. O problema de otimização consiste em encontrar o portfólio que fornece a variância mínima, fixados valores para o retorno esperado e para a assimetria. Já o problema dual consiste em maximizar a assimetria, fixados valores para o retorno esperado e para a variância.

---

[3] A diferença entre o retorno dos ativos com risco e o retorno do ativo sem risco.

Fixados, respectivamente, o retorno esperado excedente do portfólio em $E(r_p)$ e a assimetria do portfólio em σ_{p^3}, tem-se o seguinte problema de otimização: minimizar a variância sujeita à seguintes restrições retorno esperado excedente igual a $E(r_p)$ e assimetria igual a σ_{p^3} . Isto é:

$$\min_{\alpha} \sigma_{p^2} = \alpha^t M_2 \alpha \,;$$
$$M_1 + (1 - \alpha^t[1])r_f) = E(r_p);$$
$$\alpha^t M_3 \,(\alpha \otimes \alpha) = \sigma_{p^3};$$

onde M_1, M_2 e M_3 são as matrizes que contêm, respectivamente, o retorno médio, a covariância e a assimetria dos n ativos de risco, $[1]$ é um vetor $n \times 1$ cujos elementos são iguais a 1 e r_f é o retorno do ativo sem risco.

O portfólio de variância mínima é encontrado minimizando o Lagrangiano:

$$\min_{\alpha}\left(\alpha^t M_2 \alpha + \lambda_1\left[E(r_p) - \left(\alpha^t M_1 + (1 - \alpha^t[1])r_f\right)\right] + \lambda_2\left[\sigma_{p^3} - \alpha^t M_3(\alpha \otimes \alpha)\right]\right). \qquad (2.3)$$

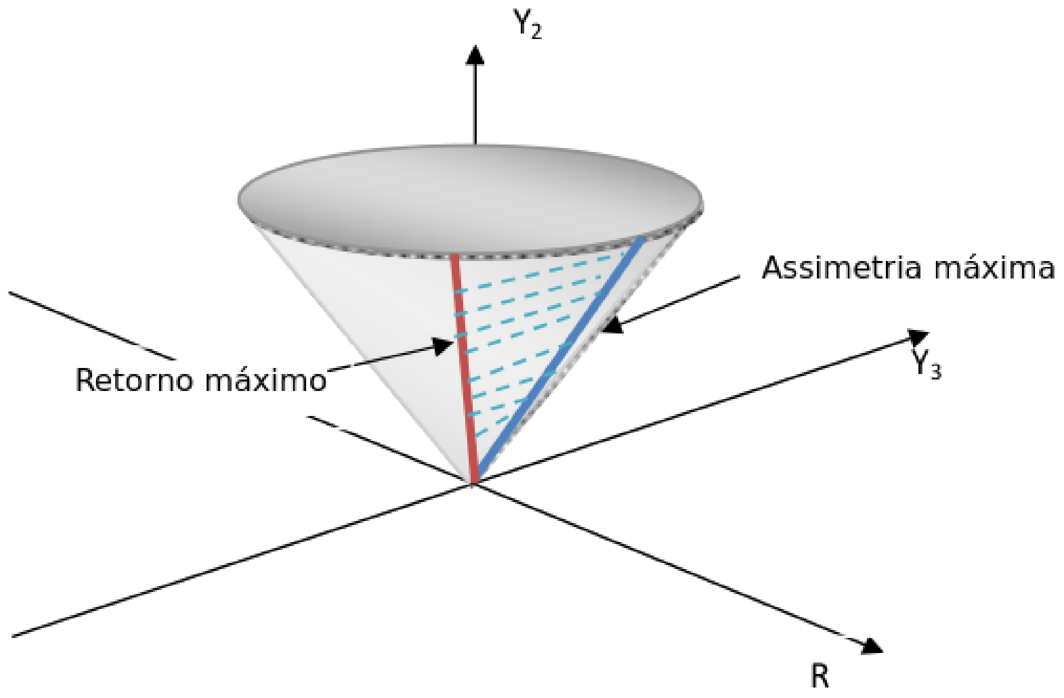
A fronteira eficiente é o conjunto de valores ótimos (soluções do problema de otimização) para diferentes valores fixados de retorno e de assimetria, pois o que é pretendido quando se constrói a fronteira eficiente é encontrar qual portfólio, dentro dessa fronteira, tem um mínimo de variância dado um nível de retorno e assimetrias fixados.

Na Figura 4 está ilustrada a fronteira eficiente (a região compreendida entre as duas linhas na qual o retorno e assimetria são respectivamente máximos) que contém triplas ordenadas, formadas pelas variâncias mínimas, e os correspondentes valores de retorno e a assimetria, os quais foram fixados.

2.2.3 Assimetria

Estudos sobre o comportamento dos investidores verificaram que, de um modo geral, os investidores buscam assimetria positiva na distribuição do retorno tolerando que a média do retorno venha a sofrer uma diminuição. Desta maneira, o investidor considera pagar mais para obter assimetria positiva no seu investimento, por ter uma diminuição do risco e um aumento da possibilidade de extremos ganhos.

Figura 4 – Representação da fronteira eficiente no espaço média-variância-assimetria.



Legenda: A fronteira eficiente é a área compreendida entre o retorno máximo e a assimetria máxima.
Fonte: Retirado de Nielsen, 2008.

A assimetria pode ser incluída como um terceiro parâmetro na teoria de análise de portfólio. Isso significa que a fronteira eficiente será estabelecida em um espaço com três dimensões ou coordenadas, isto é, média, variância e assimetria. No entanto, ao admitir o momento de terceira ordem como um parâmetro, algumas precauções precisam ser tomadas, pois a assimetria é uma medida que não possui tanta estabilidade quanto a média e a variância.

Quando o momento de terceira ordem é adotado em um modelo de análise de portfólio de ativos financeiros, é preciso ter cautela com determinados fatores (NIELSEN, 2008), como:

- Diferença de intervalos de tempo: anual, semestral, trimestral, mensal etc;
- Se o método utilizado para gerar o retorno $\bar{R}_i$ for estável e estacionário, então a medida de assimetria tende a ser mais estável. Entretanto, caso o processo adotado não seja estacionário, a medida de assimetria será errática e altamente sensível à amostra;

- O retorno pode ser calculado com função logarítmica ou a partir de aritmética simples. Quando o retorno é logarítmico a assimetria é menor que a assimetria obtida com um retorno determinado via simples aritmética;
- Verificar a estabilidade da assimetria no ponto de início do período de observação, pois a estabilidade varia de acordo com a data de início escolhida.

3 FUNÇÃO UTILIDADE E OS MOMENTOS DE ORDEM PAR E ÍMPAR

Scott & Horvath (1980) analisaram a função utilidade e os momentos centrais e verificaram que, de acordo com o crescimento ou decrescimento desses momentos, a função utilidade é maximizada. Também demonstraram que as direções de preferência para os momentos centrais de ordem par são preferencialmente negativas, entretanto, opostamente, para os momentos centrais de ordem ímpar, elas são preferencialmente positivas. Serão discutidos neste trabalho os resultados de Scott & Horvath (1980) e sua importância para a maximização da função utilidade.

Posteriormente, a partir da análise da função utilidade será construído um paralelo entre o princípio de maximização da função utilidade e os modelos de seleção de portfólio propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004).

3.1 Direções de preferência em momentos de ordem superior

O investidor busca maximizar sua função utilidade; ou seja, para otimizar a função é preciso analisar os termos que a compõem. Será visto que essa função possui momentos de ordem ímpares e pares, e de acordo com a preferência do investidor, irá optar por aumentar ou reduzir esses termos. Para tanto, serão analisados os critérios que conduzem a opção do investidor.

Sejam W a riqueza inicial do investidor e X o ganho de seus investimentos. Note que X é uma variável aleatória. A taxa de retorno de seus investimentos é dada por $r = \frac{X}{W}$. Supondo que a função utilidade do investidor dependa apenas de sua riqueza e dos ganhos de seus investimentos, considere $\mu = E(rW + W)$ o valor esperado de seus investimentos.

Denotando por U a função utilidade do investidor e expandindo a função U em série de Taylor em torno de μ, segue que:

$$U = U(\mu) + U'(\mu)(rW + W \ - \mu) + \cdots + \frac{U^{(k)}(\mu)}{k!}(rW + W \ - \mu)^k + \cdots .$$

Truncando a série no termo de ordem k e calculando o valor esperado, temos que:

$$E(U) \approx U(\mu) + U'(\mu)E[(rW + W - \mu)] + \cdots + \frac{U^{(k)}(\mu)}{k!}E[(rW + W - \mu)^k]. \quad (3.1)$$

Pelas propriedades dos momentos, $E[(rW + W - \mu)] = E(rW + W) - E(\mu) = \mu - \mu = 0$, o termo no valor esperado da função utilidade que equivale ao momento de primeira ordem é sempre igual a zero. Portanto, a estimativa do valor esperado da função utilidade é obtida pela expressão:

$$E[U] \approx U(\mu) + \frac{U''(\mu)}{2}\sigma^2 + \sum_{i=3}^{k}\frac{U^{(i)}(\mu)}{i!}\mu_i, \quad (3.2)$$

onde $U^{(i)}$ representa a i-ésima derivada de U, μ_i é o i-ésimo momento central e $\sigma^2 = \mu_2$.

Caso o valor esperado da função utilidade seja representado por uma distribuição normal, os momentos de primeira e de segunda ordem serão suficientes para caracterizar esse tipo de distribuição. Porém, caso a distribuição dos retornos do portfólio não seja normal, serão necessários momentos de ordem superior, como a assimetria, para analisar o comportamento da função utilidade. Scott & Horvath (1980) apontam que, para momentos de ordem superior a dois, duas perguntas devem ser respondidas: a direção de preferência de cada momento[4] pode ser previamente determinada? Se sim, qual é a direção de preferência de cada momento?

O princípio da não satisfação afirma que a utilidade de $X + 1$ reais será sempre maior que a utilidade de X reais, ou seja, o investidor tem preferência por ganhar sempre mais do que menos. Logo, esse princípio requer que a primeira derivada da função utilidade seja positiva, isto é, $U' > 0$. Com isso, é estabelecido que a função utilidade é crescente.

A preferência do investidor pelo risco é discutida usualmente em termos de um jogo justo, no qual o valor esperado é exatamente igual ao seu custo (CASTRO & FARO, 2005).

Definição 3.1.1. *Seja a preferência de um indivíduo representável pela função utilidade U, diz-se que o indivíduo é:*

[4] A direção de preferência do momento μ_i corresponde ao sinal de seu coeficiente na expressão $U(\mu) + \frac{U''(\mu)}{2}\sigma^2 + \sum_{i=3}^{\infty}\frac{U^{(i)}(\mu)}{i!}\mu_i$.

(a) *avesso ao risco se preferir não participar de jogos justos;*
(b) *neutro ao risco se for indiferente entre participar ou não de jogos justos;*
(c) *propenso ao risco se preferir participar de jogos justos.*

Pela aversão ao risco, o investidor prefere rejeitar o risco de um jogo justo, mantendo um retorno seguro ao invés de se submeter ao risco desse jogo. Considerando a hipótese de aversão ao risco do investidor, tal fato requer que a segunda derivada da função utilidade seja menor que zero, isto é, $U'' < 0$.

Diz-se que o investidor possui uma direção de preferência estrita pelo i-ésimo momento central, quando a i-ésima derivada $U^{(i)}(w)$ da função utilidade é positiva para todo w ou negativa para todo w ou identicamente nula. Porém, se o investidor não possuir uma direção de preferência estrita, temos $U^{(i)}(w) \geq 0$ ou $U^{(i)}(w) \leq 0,\ \forall\, w$.

O Teorema 3.1.1. estabelece que $U'''(w) > 0$ se a função utilidade tem primeira derivada positiva $U'(w) > 0$ e segunda derivada negativa $U''(w) < 0$. Note que, se a utilidade marginal da riqueza é positiva para todos os níveis de riqueza, temos $U'(w) > 0$. Além disso, se o investidor apresenta aversão consistente ao risco, temos $U''(w) < 0$.

Teorema 3.1.1. *Considere um investidor que possui utilidade marginal da riqueza positiva para todos os níveis de riqueza, que apresenta aversão consistente ao risco em todos os níveis de riqueza e tem preferência estrita pelas direções dos momentos centrais. Então, ele terá direção de preferência positiva pela assimetria.*

Demonstração: Como a preferência do investidor por momentos é estrita, teremos, para todo w, apenas uma das seguintes possibilidades: $\mathrm{U}'''(\mathrm{w}) < 0, U'''(w) = 0$ ou $\mathrm{U}'''(\mathrm{w}) > 0$. Suponha por absurdo que, para todo w, tenhamos $\mathrm{U}'''(\mathrm{w}) < 0$ (ou $\mathrm{U}'''(\mathrm{w}) = 0$). Tome uma constante real w_1 e considere $\mathrm{w}^* = \mathrm{w}_1 - \frac{\mathrm{U}'(\mathrm{w}_1)}{\mathrm{U}''(\mathrm{w}_1)}$. Observe que $\mathrm{w}_1 < \mathrm{w}^*$. Considere w_2 tal que $\mathrm{w}_2 \geq \mathrm{w}^*$. Pelo teorema do valor médio existe $\overline{\mathrm{w}} \in (\mathrm{w}_1, \mathrm{w}_2)$, tal que:

$$\mathrm{U}'(\mathrm{w}_2) - \mathrm{U}'(\mathrm{w}_1) = \mathrm{U}''(\overline{\mathrm{w}})(\mathrm{w}_2 - \mathrm{w}_1).$$

Logo, $\mathrm{U}'(\mathrm{w}_2) = \mathrm{U}'(\mathrm{w}_1) + \mathrm{U}''(\overline{\mathrm{w}})(\mathrm{w}_2 - \mathrm{w}_1)$. A hipótese $\mathrm{U}'''(\mathrm{w}) < 0$ ou $\mathrm{U}'''(\mathrm{w}) = 0, \forall\, \mathrm{w}$ diz que $\mathrm{U}''(\overline{\mathrm{w}}) \leq \mathrm{U}''(\mathrm{w}_1)$ e, consequentemente,

$$\mathrm{U}'(\mathrm{w}_2) \leq \mathrm{U}'(\mathrm{w}_1) + \mathrm{U}''(\mathrm{w}_1)(\mathrm{w}_2 - \mathrm{w}_1).$$

Segue de $w_2 \geq w^*$ que $w_2 - w_1 \geq -\frac{U'(w_1)}{U''(w_1)}$. Como $U''(w) < 0$, para todo w, temos:

$$U'(w_2) \leq U'(w_1) + U''(w_1)(w_2 - w_1) \leq U'(w_1) + U''(w_1)\left(-\frac{U'(w_1)}{U''(w_1)}\right) = 0,$$

em contradição com $U'(w) > 0, \forall\, w$. Conclui-se, então, que $U'''(w) > 0, \forall\, w$.

■

Se a terceira derivada é positiva, então o investidor tem preferência por assimetria positiva na distribuição do retorno do portfólio. O investidor possui aversão ao risco, portanto, o ideal é que a variância seja a menor possível. Se a variância for baixa, a distribuição da função utilidade será mais achatada que a distribuição normal, sendo caracterizada como platicúrtica.

Teorema 3.1.2. *A aversão consistente ao risco, a preferência estrita pelas direções dos momentos centrais e a preferência para que a assimetria seja positiva implicam em direção de preferência negativa para o quarto momento central (curtose). Isto é, se para todo w temos $U''(w) < 0$ e $U'''(w) > 0$ e há preferência estrita pelas direções dos momentos centrais, então $U^{(4)}(w) < 0,\ \forall\, w$.*

Demonstração: Como a preferência do investidor por momentos é estrita, teremos, para todo w, apenas uma das seguintes possibilidades: $U^{(4)}(w) < 0, U^{(4)}(w) = 0$ ou $U^{(4)}(w) > 0$. Suponha por absurdo que $U^{(4)}(w) > 0$ ou $U^{(4)}(w) = 0, \forall$ w. Tome uma constante real w_1 e considere $w^* = w_1 - \frac{U''(w_1)}{U'''(w_1)}$. Observe que $w_1 < w^*$. Considere w_2 tal que $w_2 \geq w^*$. Pelo teorema do valor médio existe $\overline{w} \in (w_1, w_2)$, tal que:

$$U''(w_2) - U''(w_1) = U'''(\overline{w})(w_2 - w_1).$$

Logo, $U''(w_2) = U''(w_1) + U'''(\overline{w})(w_2 - w_1)$. A hipótese $U^{(4)}(w) > 0$ ou $U^{(4)}(w) = 0, \forall$ w diz que $U'''(\overline{w}) \geq U'''(w_1)$ e, consequentemente,

$$U''(w_2) \geq U''(w_1) + U'''(w_1)(w_2 - w_1).$$

Segue de $w_2 \geq w^*$ que $w_2 - w_1 \geq -\frac{U'(w_1)}{U''(w_1)}$. Como $U'''(w) > 0$, para todo w, temos:

$$U''(w_2) \geq U''(w_1) + U'''(w_1)(w_2 - w_1) \geq U''(w_1) + U'''(w_1)\left(-\frac{U'(w_1)}{U''(w_1)}\right) = 0,$$

em contradição com $U''(w) < 0, \forall\, w$. Conclui-se, então, que $U^{(4)}(w) < 0, \forall\, w$.

■

Podem ser feitas análises semelhantes para as derivadas de ordem mais alta, onde percebe-se a preferência positiva do investidor pelos momentos centrais de ordem ímpar e negativa para os momentos centrais de ordem par.

Muitas evidências mostram que a aversão ao risco do investidor a partir de um determinado momento começa a decair. Isso se dá devido à satisfação do investidor com o aumento dos seus investimentos ao longo do tempo a partir do acúmulo de capital. Sendo assim, o investidor começa a tolerar riscos maiores nos seus investimentos. Pode-se dizer que, quanto maior a riqueza acumulada, maior será a tolerância ao risco. De mesmo modo, no sentido inverso, quanto menor a riqueza acumulada, menor será a tolerância ao risco.

O investidor tem preferência por assimetria positiva em consequência de a probabilidade de ganhos extremos ser maior e a perda ser limitada. Tome o caso das pessoas que compram tíquete de loteria: o retorno prometido pelo prêmio da loteria é extremamente alto, caracterizando uma assimetria positiva, em contrapartida o risco envolvido nesse tipo de investimento é limitado ao valor do tíquete.

De modo inverso, as pessoas possuem aversão à assimetria negativa. Considere o exemplo de uma pessoa que compra um seguro: esse indivíduo considera válido pagar a companhia de seguros um valor que ele não irá recuperar posteriormente com o intuito de evitar grandes perdas financeiras, por isso as pessoas fazem seguros de carro, casa etc. Essa extrema perda financeira seria caracterizada como uma assimetria negativa no retorno do investidor.

3.2 **Maximização da Função Utilidade**

Por meio de uma abordagem axiomática, pode-se estabelecer que escolhas racionais do investidor, em situações envolvendo incertezas, podem ser representadas por uma função utilidade (NEUMANN & MORGENSTERN, 2004). Na análise de um portfólio de ativos pelo método proposto por Markowitz, não se consideram momentos de ordem superior da função utilidade, todos os momentos com ordem maior que dois são considerados nulos. Com isso, é possível assumir, a partir da equação (3.2), que a função utilidade esperada $E_M(U)$ é dada por:

$$E_M(U) = U(\mu_1) + \frac{U''(\mu_1)}{2}\mu_2.$$

Como dito anteriormente, caso a distribuição presente na função utilidade seja normal, o primeiro e segundo momentos centrais são suficientes para descrever seu comportamento. Entretanto, se a distribuição de retorno dos ativos não for simétrica, será necessário o uso do terceiro momento central para estudar a função utilidade do investidor. Em suma, o modelo proposto por Athayde e Flôres terá um resultado mais consistente na aproximação da função utilidade.

Similarmente, no modelo de Athayde & Flôres Jr. (2004), pode-se assumir, a partir da equação (3.2), que a função utilidade esperada $E_F(U)$ é dada por:

$$E_F(U) = U(\mu_1) + \frac{U''(\mu_1)}{2}\mu_2 + \frac{U'''(\mu_1)}{6}\mu_3.$$

Para estudar o comportamento da função utilidade esperada, será assumido que a utilidade marginal é positiva, a aversão absoluta ao risco em todos os níveis de riqueza e preferência estrita consistente relativa aos momentos centrais e que o princípio da não satisfação é válido.

Teorema 3.2.1. *Seja a função utilidade esperada $E_M(U)$ dada por:*

$$E_M(U) = U(\mu_1) + \frac{U''(\mu_1)}{2}\mu_2.$$

Para maximizar $E_M(U)$ com um dado retorno esperado (μ_1 é fixado), é necessário e suficiente minimizar a variância μ_2.

Demonstração: Considere que um determinado valor para μ_1 fixo e $U''(\mu_1) < 0$ (a aversão ao risco é decrescente) para maximizar $E_M(U)$ é necessário e suficiente minimizar a variância μ_2.

■

O objetivo do investidor é sempre aumentar o retorno de seus investimentos. Fica provado que o investidor tem preferência positiva para os momentos de ordem ímpar na função utilidade e,de forma oposta, o investidor tem preferência negativa para os momentos de ordem par (SCOTT & HORVATH, 1980). Isso se dá devido aos momentos centrais de ordem par representarem a volatilidade presente no portfólio.

Teorema 3.2.2. *Seja a função utilidade esperada $E_F(U)$ dada por:*

$$E_F(U) = U(\mu_1) + \frac{U''(\mu_1)}{2}\mu_2 + \frac{U'''(\mu_1)}{6}\mu_3.$$

Então,

1. *para maximizar $E_F(U)$ fixados os valores do retorno esperado e da assimetria (μ_1 e μ_3 são fixados), é necessário e suficiente minimizar a variância μ_2;*
2. *para maximizar $E_F(U)$ fixados os valores do retorno esperado e da variância (μ_1 e μ_2 são fixados), é necessário e suficiente maximizar a assimetria μ_3.*

Demonstração:

1. Considerando μ_1 e μ_3 fixados, para maximizar $E_F(U)$ é necessário e suficiente maximizar $\frac{U''(\mu_1)}{2}\mu_2$. O decrescimento da aversão absoluta ao risco implica que $U''(\mu_1) < 0$. Portanto, o máximo é alcançado se, e somente se, minimizarmos a variância μ_2.
2. Assumindo que μ_1 e μ_2 são fixados, para maximizar $E_F(U)$ é necessário e suficiente maximizar $\frac{U'''(\mu_1)}{6}\mu_3$. Considerando o Teorema 3.1.1, temos $U'''(\mu_1) > 0$. Portanto, o máximo é alcançado se, e somente se, maximizarmos a assimetria μ_3.

■

Foram utilizados os resultados de Scott & Horvath (1980) para relacionar o problema de otimização proposto por Markowitz (1952) e Athayde & Flôres Jr. (2004) e o princípio de maximização do valor esperado da função utilidade do investidor, quando se admite que ela é dada pela equação (3.1). Os resultados obtidos são consistentes com o comportamento esperado do investidor de sempre minimizar o risco envolvido nos ativos financeiros em que investe. Porém, de um modo geral, os ativos com um risco baixo apresentam pouca possibilidade de grandes retornos. Com isso, o investidor passa a ter de buscar um equilíbrio entre o retorno esperado e o risco que está disposto a correr.

No caso em que o investidor fixe o retorno esperado e também fixe a assimetria presente no portfólio, isto é, fixe os momentos centrais de ordem um e três, de tal modo que apenas o momento de ordem dois varie, para maximizar a função utilidade o investidor irá optar por minimizar o momento de ordem dois, ou seja, minimizar o risco presente no portfólio. Ressalta-se que, ao fixar a assimetria, a preferência do investidor continua sendo por uma assimetria positiva. Logo, nesse caso, o investidor busca minimizar o risco, considerando um determinado retorno médio e assimetria fixos. No caso em que o investidor apresente preferência por manter um determinado retorno e risco fixados, a maximização do valor esperado da utilidade corresponde a maximizar o momento central de ordem três, a assimetria. Esse resultado dialoga com os resultados de dualidade provados por Athayde & Flôres Jr. (2004). Isto é, o problema de minimizar a variância, quando estão fixados o retorno e a assimetria, e o problema de maximizar a assimetria, quando estão fixados o retorno e a variância, são versões do princípio de maximização do valor esperado da função utilidade esperada do investidor.

De acordo com a preferência do investidor, os três primeiros momentos centrais podem variar. A partir disso surge a importância de estabelecer uma análise sobre esses momentos de modo a aumentar a precisão da aproximação da função utilidade, visando satisfazer o investidor de acordo com a vontade dele de sempre aumentar seus lucros.

4 DIVERSIFICAR COMPENSA

A diversificação é extremamente usada por economistas na formação de uma carteira de investimentos, pois ela possibilita manter o retorno esperado e diminuir o risco presente em um portfólio. Samuelson (1967) demonstrou a importância da diversificação em uma carteira de investimento provando os resultados esperados da diversificação em um portfólio. Será discutida detalhadamente a análise feita por Samuelson (1967).

4.1 Compensa diversificar – uma demonstração

O provérbio “não coloque todos os ovos na mesma cesta” é bastante popular. Os economistas, como Markowitz, que trabalham apenas com a renda média e sua variância, traduzem essa regra para o universo das carteiras de investimento da seguinte maneira: ao investir um total fixo da riqueza em títulos independentes e identicamente distribuídos, este total deve ser igualmente distribuído entre os diferentes títulos. Deste modo, o retorno médio permanece inalterado e a variância presente na carteira composta por esses investimentos será minimizada, diminuindo, assim, a volatilidade presente no portfólio de ativos financeiros.

Antes de demonstrar a importância da diversificação em um portfólio de ativos financeiros, é preciso definir alguns conceitos que serão assumidos para a função $U(X)$ e a regra da cadeia.

Definição 4.1.1. *A função $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ é dita suave ou de classe C^∞ $(U, \mathbb{R}^n)$ se for de classe $C^n(U, \mathbb{R}^n)$ para todo n.*

Definição 4.1.2. *Seja a matriz $M_{n\,x\,n}$ e a forma quadrática associada à matriz dada pela função escalar $Q: \mathbb{R}^n \to \mathbb{R}$, na qual $Q(h) = h^T \cdot M \cdot h$, temos:*

$$\mathrm{Q}(\mathrm{h}_1, \mathrm{h}_2, \dots, \mathrm{h}_\mathrm{n}) = \begin{bmatrix} h_1 & \dots & h_n \end{bmatrix} \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix},$$

ou ainda,

$$Q(h_1, h_2, \dots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j \,.$$

A matriz M é dita negativa semi-definida se $Q(h) = h^T \cdot M \cdot h \leq 0$. Se a desigualdade é estrita para todo $h \neq 0$, então M é dita negativa definida.

Definição 4.1.3. *A função $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ de classe C^2 é côncava se, e somente se,*

$$f\big((1 - t) \cdot p + t \cdot q\big) \geq (1 - t) \cdot f(p) + t \cdot f(q) \,,$$

para todo $p, q \in U$ e todo $t \in [0, 1]$. Se a inequação é estrita para todo $p, q \in U$ e todo $t \in [0, 1]$, então a função f é estritamente côncava.

Teorema 4.1.1. *Considere a função $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ de classe C^2 e p um ponto crítico de f no interior do conjunto U. Se $U^2 f(p)$ é uma matriz negativa definida, então p é um ponto de máximo local de f.*

Definição 4.1.4. *Uma função $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ é dita simétrica quando não é alterada por quaisquer permutações de suas variáveis.*

Teorema 4.1.2. (Regra da cadeia) *Considere $f: D_f \subset \mathbb{R}^n \to \mathbb{R}^m$ e $g: D_g \subset \mathbb{R}^m \to \mathbb{R}^k$ duas funções de classe C^k com $k \geq 1$, tais que $f\big(D_f\big) \subset D_g$ de modo que seja possível construir a composição $h = g \circ f: D_h \subset \mathbb{R}^n \to \mathbb{R}^k$. Então, h também é uma função de classe C^k e a matriz jacobiana de g em $f(p)$ pela matriz jacobiana de f em p é dada por:*

$$Dh(p) = D(g \circ f)(p) = Dg\big(f(p)\big) \cdot Df(p) \,.$$

Teorema 4.1.3. *Seja $U(X)$ uma função suave e estritamente côncava, que é monótona para valores de X não negativos, e sejam $X_1, \dots, X_n$ variáveis aleatórias independentes, identicamente distribuídas com frequência de distribuição conjunta, temos:*

$$P\{X_1 \leq a_1, X_2 \leq a_2, \dots, X_n \leq a_n\} = F(a_1)F(a_2)\cdots F(a_n).$$

Considere a função ψ dada por:

$$\psi(\lambda_1, \dots, \lambda_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} U\left(\sum_{j=1}^{n} \lambda_j x_j\right) dF(x_1) \cdots dF(x_n)\,.$$

Suponha, ainda, que as variáveis tenham mesma média e variância (finita). Isto é,

$$E[X_i] = \int_{-\infty}^{\infty} x_i \, dF(x_i) = \mu_1$$

e

$$E[X_i - \mu_i]^2 = \int_{-\infty}^{\infty} (x_i - \mu_1)^2 \, dF(x_i) = \mu_2$$

com $0 < \mu_2 < \infty$. Então, $\psi(\lambda_1, \dots, \lambda_n)$ é uma função simétrica estritamente côncava e coincide com $E\left[U\left(\sum_{j=1}^{n} \lambda_j X_j\right)\right]$. Além disso, o máximo de $\psi(\lambda_1, \dots, \lambda_n)$ sujeito a $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$ e $\lambda_i \geq 0$ é assumido em $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$.

Demonstração: Seja $\mathrm{h}(\mathrm{x}_1, \dots, \mathrm{x}_n) = \sum_{j=1}^{n} \lambda_j \mathrm{x}_j$. Como $\mathrm{U}(\mathrm{X})$ é contínua, a função $\mathrm{g} = \mathrm{U} \circ \mathrm{h}$ é contínua. Logo, g é Borel mensurável. Portanto, o Corolário 1.1.1. diz que:

$$\mathrm{E}\left[\mathrm{U}\left(\sum_{j=1}^{n} \lambda_j \mathrm{X}_j\right)\right] = \int_{\mathbb{R}^n} \mathrm{g}(\mathrm{x}) \, \mathrm{dF}(\mathrm{x})\,.$$

Como $\mathrm{X}_1, \dots, \mathrm{X}_n$ são variáveis aleatórias independentes e identicamente distribuídas, temos $\mathrm{F}(\mathrm{x}) = \mathrm{F}(\mathrm{x}_1)\mathrm{F}(\mathrm{x}_2)\cdots\ \mathrm{F}(\mathrm{x}_n)$. Logo,

$$\mathrm{E}\left[\mathrm{U}\left(\sum_{j=1}^{n} \lambda_j \mathrm{X}_j\right)\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathrm{U}\left(\sum_{j=1}^{n} \lambda_j \mathrm{x}_j\right) \mathrm{dF}(\mathrm{x}_1) \cdots \mathrm{dF}(\mathrm{x}_n) = \psi(\lambda_1, \dots, \lambda_n).$$

Para provar que $\psi(\lambda_1, \dots, \lambda_n)$ é simétrica, é preciso mostrar que ela é invariante em relação a permutações dos λ_i. Como as variáveis X_i são identicamente distribuídas, cada uma das integrações é realizada em relação à mesma medida (a associada à distribuição F). Além disso, as integrações são realizadas sobre toda a reta. Logo, através de mudanças de variável e da aplicação do teorema de Fubini, conclui-se que $\psi(\lambda_1, \dots, \lambda_n)$ é simétrica.

Considere

$$\lambda^1 = (\lambda_1^1, \dots, \lambda_n^1), \qquad \lambda^2 = (\lambda_1^2, \dots, \lambda_n^2)$$

e $\alpha \in (0,1)$. Denote $W = (X_1, \dots, X_n)$. A concavidade de U garante que:

$$\begin{aligned} \psi(\alpha\lambda^1 + (1-\alpha)\lambda^2) &= \mathrm{E}\big[\mathrm{U}\big((\alpha\lambda^1 + (1-\alpha)\lambda^2) \cdot \mathrm{W}\big)\big] \\ &> E[\alpha\, \mathrm{U}(\lambda^1 \cdot \mathrm{W}) + (1-\alpha)\mathrm{U}(\lambda^2 \cdot \mathrm{W})] \\ &= \alpha\, \mathrm{E}[\mathrm{U}(\lambda^1 \cdot \mathrm{W})] + (1-\alpha)\mathrm{E}[\mathrm{U}(\lambda^2 \cdot \mathrm{W})] \\ &= \alpha\psi(\lambda^1) + (1-\alpha)\psi(\lambda^2). \end{aligned}$$

Consequentemente, ψ é estritamente côncava.

Para derivar parcialmente a função ψ com respeito a λ_i, deriva-se sob o sinal de integração e obtém-se:

$$\frac{\partial\psi}{\partial\lambda_i}(\lambda_1, \dots, \lambda_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i\, \mathrm{U}'\left(\sum_{k=1}^{n} \lambda_k\, x_k\right) \mathrm{d}F(x_1) \cdots \mathrm{d}F(x_n). \tag{4.1}$$

Associa-se ao problema de otimização da função ψ com a restrição $\sum_{j=1}^{n} \lambda_j = 1$, o seguinte lagrangiano:

$$L(\lambda_1, \dots, \lambda_n, \alpha) = \psi(\lambda_1, \dots, \lambda_n) + [(\lambda_1 + \cdots + \lambda_n) - 1]\alpha.$$

Como a função U é suave, existem todas as derivadas parciais de ψ. Logo, pela condição de primeira ordem,

$$\begin{cases} \dfrac{\partial L}{\partial \lambda_1} = \dfrac{\partial \psi}{\partial \lambda_1} + \alpha = 0 \\ \qquad\quad \vdots \\ \dfrac{\partial L}{\partial \lambda_n} = \dfrac{\partial \psi}{\partial \lambda_n} + \alpha = 0 \end{cases} .$$

Portanto, teremos $\frac{\partial \psi}{\partial \lambda_1} = \frac{\partial \psi}{\partial \lambda_2} = \cdots = \frac{\partial \psi}{\partial \lambda_n}$.

Observe que cada uma das integrações em (4.1) é realizada em toda a reta e, como as variáveis são identicamente distribuídas, são todas realizadas com relação à mesma medida. Além disso, para $(\lambda_1, \dots, \lambda_n) = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$, os coeficientes de x_k em $\sum_{k=1}^{n} \lambda_k\, x_k$ são idênticos. Portanto, a derivada $\frac{\partial \psi}{\partial \lambda_i}$ independe de i e teremos:

$$\frac{\partial \psi}{\partial \lambda_1}\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = \cdots = \frac{\partial \psi}{\partial \lambda_n}\left(\frac{1}{n}, \dots, \frac{1}{n}\right).$$

Derivando novamente com relação a λ_j, obtém-se os elementos da matriz Hessiana de ψ,

$$\frac{\partial^2 \psi}{\partial \lambda_i\, \partial \lambda_j} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i\, x_j\, U''\left(\sum_{k=1}^{n} \lambda_k\, x_k\right) dF(x_1) \cdots dF(x_n).$$

Por conseguinte, devido à hessiana ser definida como negativa, isto é, a forma quadrática da hessiana $\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 \psi}{\partial \lambda_i \lambda_j} y_i y_j < 0$, pela condição de segunda ordem temos o máximo de ψ quando $\lambda_1 = \lambda_2 = \cdots = \lambda_n = \frac{1}{n}$ e esse máximo é único (pela concavidade estrita e a restrição afim).

■

4.1.1 <u>O caso geral de interdependência simétrica</u>

No Teorema 4.1.3., supõe-se que as variáveis aleatórias $X_1, \ldots, X_n$ eram independentes e adota-se que sua distribuição conjunta era dada por $F(x_1) \cdots F(x_n)$. Agora, substituindo essa suposição pela hipótese da distribuição simétrica, será assumido que:

$$P\{X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n\} = F(a_1, a_2, \ldots, a_n),$$

onde F é uma função simétrica. É possível desconsiderar, pela sua trivialidade, o caso em que os $x's$ estão conectados por uma relação funcional exata, que em vista da simetria é necessariamente da forma

$$x_1 = x_2 = \cdots = x_n = x$$

e

$$P\{X \leq x\} = F(x, \ldots, x).$$

Assume-se que são finitas as médias, variâncias e covariâncias. Além disso, supõe-se que são as mesmas para todas as variáveis. Isto é,

$$E[X_i] = \int_{-\infty}^{\infty} x_i \, dF(x_1, \ldots, x_n) = \mu \,,$$
$$E[(X_i - \mu_i)^2] = \int_{-\infty}^{\infty} (x_i - \mu_i)^2 \, dF(x_1, \ldots, x_n) = \mu_2$$

e

$$E\left[(X_i - \mu)\left(X_j - \mu\right)\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_i - \mu)\left(x_j - \mu\right) dF(x_1, \ldots, x_n) \; = \sigma_{ij}.$$

Observe que as variâncias e as covariâncias calculadas acima são elementos de uma matriz semi-definida positiva. Por meio do que foi estabelecido, é possível agora generalizar a teoria sobre a diversificação.

Teorema 4.1.4. *Seja $U(X)$ uma função suave, estritamente côncava. Considere a função ψ dada por:*

$$\phi(\lambda_1, \dots, \lambda_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} U\left(\sum_{j=1}^{n} \lambda_j\, x_j\right) dF(x_1, \dots, x_n)\,.$$

Então, $\phi(\lambda_1, \dots, \lambda_n)$ é uma função simétrica e estritamente côncava e coincide com $E\left[U\left(\sum_{i=1}^{n} \lambda_j\, x_j\right)\right]$. Além disso, o máximo de $\phi(\lambda_1, \dots, \lambda_n)$ sujeito a $\sum_{i=1}^{n} \lambda_i = 1$, com $\lambda_i \geq 0$, é assumido em $\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$.

O teorema diz que sempre vale a pena diversificar!

Demonstração: A demonstração é análoga à do Teorema 4.1.3.. Segue novamente da continuidade e concavidade estrita de U, que $\phi(\lambda_1, \dots, \lambda_n)$ é estritamente côncava. Analogamente, usando mudança de variável nas integrais e explorando a simetria da distribuição F, deduz-se a simetria de ϕ.

A condição de primeira ordem do problema de otimização condicionado diz que é preciso ter

$$\frac{\partial \phi}{\partial \lambda_1}(\lambda_1, \dots, \lambda_n) = \cdots = \frac{\partial \phi}{\partial \lambda_n}(\lambda_1, \dots, \lambda_n).$$

A simetria de F diz que o valor da derivada parcial de ϕ com respeito a λ_i calculada em $(\lambda_1, \dots, \lambda_n) = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ independe de i. Logo, temos:

$$\frac{\partial \phi}{\partial \lambda_1}\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = \cdots = \frac{\partial \phi}{\partial \lambda_n}\left(\frac{1}{n}, \dots, \frac{1}{n}\right).$$

Os elementos da matriz Hessiana ϕ são dados por:

$$\frac{\partial^2 \phi}{\partial \lambda_i\, \partial \lambda_j}(\lambda_1, \dots, \lambda_n) \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots\; x_i\, x_j\, U''\left(\sum_{i=k}^{n} \lambda_k\, x_k\right) dP(x_1, \dots, x_n)\,,$$

que é uma matriz de Gram definida negativa. A partir disso, a concavidade de ϕ é confirmada e quando $\phi\left(\frac{1}{n},\ldots,\frac{1}{n}\right)$ tem-se o máximo da função com a restrição $\sum \lambda_i = 1$.

■

Para a diversificação corresponder à distribuição da riqueza igualmente entre os ativos, ou seja, os pesos atribuídos a cada investimento serem idênticos, então é necessário supor que a distribuição de probabilidade apresenta simetria ou alguma propriedade análoga. Para verificar esse fato, é interessante supor que as variáveis (X_1, X_2, X_3) são independentes, com X_2 e X_3 tendo a mesma distribuição $F(X_i)$, e com X_1 tendo uma distribuição que é idêntica a $\frac{1}{2}X_2 + \frac{1}{2}X_3$, isto é[5],

$$P\{X_1 \leq x_1\} = Q(x_1) = \int_{-\infty}^{\infty} F(2x_1 - 2s)\, dF(2s).$$

Nesse caso, a simetria revela que a riqueza deve ser dividida igualmente entre os investimentos X_1 e $\frac{1}{2}X_2 + \frac{1}{2}X_3$. Isto é equivalente a investir em (X_1, X_2, X_3) na fração $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$. Ao se trabalhar com dois momentos, média e variância, é evidente que a variância mínima não é obtida com $(\lambda_1, \ldots, \lambda_n) = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ quando se tem $\sigma_{ii} \neq \sigma_{jj}, \sigma_{ij} \neq \sigma_{rs}$ (SAMUELSON, 1967).

É possível provar que alguma diversificação é imperativa em circunstâncias bastante gerais. Considere os investimentos $(x_1, \ldots, x_n)$. Assuma que todos eles têm a mesma média comum e que suas variâncias são finitas e não-nulas. Finalmente, suponha que uma das variáveis, x_1, é independente das demais. Então, a carteira ótima deve ter $\lambda_1^* > 0$. Isto é, há algum investimento em x_1, como será mostrado no Teorema 4.1.5..

Teorema 4.1.5. *Seja $(X_1, X_2, \ldots, X_n)$ um vetor de variáveis aleatórias. Suponha que sua distribuição de probabilidade seja dada por $P(x_1)Q(x_2, \ldots, x_n)$. Além disso, assuma que elas têm a mesma média e que suas variâncias sejam positivas e finitas. Isto é,*

---

[5] Para calcular a distribuição de $\boldsymbol{X_1}$, basta usar o fato de que se $\boldsymbol{X}$ é uma variável aleatória com distribuição $\boldsymbol{F_X}$, então a distribuição $\boldsymbol{F_Y}$ da variável $\boldsymbol{Y = aX}$ é dada por $\boldsymbol{F_Y(y) = F_X\left(\frac{y}{a}\right)}$ (BEAN, 2001, p. 282). Além disso, se $\boldsymbol{X}$ e $\boldsymbol{Z}$ são variáveis independentes, e $\boldsymbol{Y = X + Z}$, então a distribuição de $\boldsymbol{Y}$ é dada pela convolução entre as ditribuições de $\boldsymbol{X}$ e $\boldsymbol{Z}$ (BEAN, 2001, p. 337).

$$E[X_i] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i \, dP(x_1) dQ(x_2, \dots, x_n) = \mu,$$
$$0 < E[(X_i - \mu)^2] < \infty.$$

Considere

$$\theta(\lambda_1, \lambda_2, \dots, \lambda_n) = E\left[U\left(\sum_{i=1}^{n} \lambda_j \, X_j\right)\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots U\left(\sum_{i=1}^{n} \lambda_j \, x_j\right) \, dP(x_1) dQ(x_2, \dots, x_n),$$

tal que θ *é uma função estritamente côncava sempre que* $U'' < 0$. *Então, se* $\theta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) = \max \ \theta(\lambda_1, \lambda_2, \dots, \lambda_n)$ *sujeito à* $\sum_{j=1}^{n} \lambda_j = 1$, *com* $\lambda_i \geq 0$, *necessariamente* $\lambda_1^* > 0$ *e* $\lambda_1^* < 1$.

Demonstração: Primeiramente, será comprovado para o caso $n = 2$ e depois generalizaremos a prova. Denotando $\frac{\partial \theta(\lambda_1, \lambda_2)}{\partial \lambda_i} = \theta_i(\lambda_1, \lambda_2)$ e considerando $\lambda_1 = 0$ e $\lambda_2 = 1$, temos:

$$\begin{aligned}
\theta_1(0,1) - \theta_2\,(0,1) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 U'(x_2) \, dP(x_1) dQ(x_2) - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 U'(x_2) \, dP(x_1) dQ(x_2) \\
&= \int_{-\infty}^{\infty} x_1 \, dP(x_1) \int_{-\infty}^{\infty} U'(x_2) \, dQ(x_2) - \int_{-\infty}^{\infty} x_2 \, U'(x_2) \, dQ(x_2) \\
&= E[X_1] E[U'(X_2)] - E[X_2 \, U'(X_2)] \\
&= \mu \, E[U'(X_2)] - E[X_2 U'(X_2)] \\
&= E[X_2] E[U'(X_2)] - E[X_2 U'(X_2)] \\
&= E\big[X_2 E[U'(X_2)]\big] - E[X_2 U'(X_2)] \\
&= E[X_2 E[U'(X_2)] - X_2 U'(X_2)].
\end{aligned}$$

Somando e subtraindo $E[\mu \, U'(X_2)] = E\big[\mu \, E[U'(X_2)]\big]$, temos:

$$\begin{aligned}
&= E\big[X_2\big(E[U'(X_2)] - U'(X_2)\big) - \mu\big(E[U'(X_2)] - U'(X_2)\big)\big] \\
&= -E[(X_2 - \mu)(U'(X_2) - E[U'(X_2)])].
\end{aligned}$$

Como $U''(X_2) < 0$, a função U' é monótona decrescente. Como $E[(x_2 - \mu)(U'(x_2) - E[U'(x_2)])]$ é a covariância entre X_2 e $U'(X_2)$ e U' é monótona decrescente, temos[6]:

[6] Se $\boldsymbol{g}$ é uma função monótona decrescente e $\boldsymbol{X}$ é uma variável aleatória, então a covariância entre $\boldsymbol{X}$ e $\boldsymbol{g}(\boldsymbol{X})$ é negativa. Para uma demonstração deste fato, consulte Schmidt (2003).

$$-E[(X_2 - \mu)(U'(X_2) - E[U'(X_2)])] > 0.$$

Note que, devido à restrição $\sum_{j=1}^{n} \lambda_j = 1$, para que (λ_1, λ_2) seja extremo de $\theta(\lambda_1, \lambda_2)$, é necessário que as derivadas parciais de θ em (λ_1, λ_2) coincidam. Consequentemente, $(\lambda_1, \lambda_2) = (0,1)$ não pode ser o ponto de máximo. Logo, necessariamente temos $\lambda_1^* > 0$. Um argumento análogo mostra que $(\lambda_1, \lambda_2) = (1,0)$ não pode ser o ponto de máximo. Logo, $\lambda_1^* < 1$.

O caso $n > 2$ será reduzido ao caso anterior. Considere $(0, \lambda_2^{**}, \dots, \lambda_n^{**})$ solução do problema de otimização condicionado

$$\max_{\lambda_2, \dots, \lambda_n} \theta(0, \lambda_2, \dots, \lambda_n)$$

com $\sum_{j=2}^{n} \lambda_j = 1$.

Seja $\theta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) = \max \quad \theta(\lambda_1, \lambda_2, \dots, \lambda_n)$ sujeito à $\sum_{j=1}^{n} \lambda_j = 1$, com $\lambda_i \geq 0$. Para provar que $\lambda_1^* \neq 0$, será mostrado que:

$$\theta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) > \theta(0, \lambda_2, \dots, \lambda_n),$$

com $\sum_{j=2}^{n} \lambda_j = 1$. Para provar essa desigualdade, basta mostrar que:

$$\theta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) > \theta(0, \lambda_2^{**}, \dots, \lambda_n^{**}),$$

Seja $\lambda_1 \in [0,1]$ e considere $\lambda_{II} = 1 - \lambda_1$, $X_{II} = \sum_{j=2}^{n} \lambda_j^{**} X_j$ e

$$\tilde{\theta}(\lambda_1, \lambda_{II}) = \theta(\lambda_1, \lambda_{II}\lambda_2^{**}, \dots, \lambda_{II}\lambda_n^{**}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\lambda_1 x_1 + \lambda_{II} x_{II}) \, dP(x_1) \, d\tilde{Q}(x_{II}).$$

Pelo caso $n = 2$, sabe-se que o ponto máximo $(\widetilde{\lambda_1}, \widetilde{\lambda_{II}})$ de $\tilde{\theta}$ quando $\lambda_1 + \lambda_{II} = 1$ é tal que $0 < \widetilde{\lambda_1} < 1$. Então, temos:

$$\theta(\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*) \geq \theta(\widetilde{\lambda_1}, \widetilde{\lambda_{II}}\lambda_2^{**}, \dots, \widetilde{\lambda_{II}}\lambda_n^{**})$$

$$= \tilde{\theta}\left(\widetilde{\lambda_1}, \widetilde{\lambda_{II}}\right)$$
$$> \tilde{\theta}(0,1) = \theta(0, \lambda_2^{**}, \dots, \lambda_n^{**})$$
$$\geq \theta(0, \lambda_2, \dots, \lambda_n)$$

com $\sum_{j=2}^{n} \lambda_j = 1$.

Portanto, $\lambda_1^* > 0$. Analogamente, $\lambda_1^* < 1$.

■

Por conseguinte, utilizando o resultado do Teorema 4.1.5., é possível enunciar dois corolários que se aplicam a títulos de risco.

Corolário 4.1.1. *Se um ativo tem uma média, pelo menos, tão boa quanto a de qualquer outro da carteira, e é independente dos demais ativos, no portfólio ótimo uma fração positiva do capital é efetivamente aplicada neste ativo.*

Corolário 4.1.2. *Se todos os investimentos têm uma média comum e são independentes, no portfólio ótimo uma fração positiva do capital é efetivamente aplicada em cada um deles.*

Quando se desconsidera a suposição de independência estrita de ativos com uma média comum, não é possível mostrar que, no portfólio ótimo, uma fração positiva do capital é efetivamente aplicada em cada um deles (SAMUELSON, 1967).

Como demonstrado, realmente compensa fracionar a riqueza total entre diversos ativos financeiros. Ao diversificar, reduz-se a volatilidade e, com isso, evita-se a possibilidade de grandes perdas financeiras em um único ativo. E, concomitantemente, mantém-se o retorno esperado sobre a carteira de investimento. Deste modo, consegue-se otimizar o portfólio, obtendo o máximo retorno esperado.

Pode-se considerar também que, com a diversificação dos investimentos, torna-se possível reduzir a correlação entre os diferentes ativos, propiciando, assim, uma menor volatilidade na carteira. Porém, ainda assim, não é possível reduzir a correlação entre os ativos de modo a ter uma correlação fortemente negativa entre eles. Samuelson (1967) verificou que, quando há um conjunto de investimentos x_i com uma média comum, possuindo correlação negativa, sendo considerada a distribuição de probabilidade condicional, pode-se afirmar que todos esses investimentos devem ter frações positivas da riqueza total investidas neles no portfólio ótimo.

Conclui-se, de forma geral, que a compensação da diversificação dos ativos em uma carteira de investimento pode ser provada para modelos que são livres da suposição restritiva de considerar apenas dois momentos centrais.

CONCLUSÃO

Os modelos de seleção de carteira de investimentos propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004) auxliam na escolha do portfólio ótimo, de tal modo que a fronteira eficiente é estabelecida. Assim sendo, pode-se determinar o conjunto de frações da riqueza que será investida nos ativos financeiros, esperando conseguir um aumento da riqueza acumulada do investidor.

Com o modelo de Athayde & Flôres Jr. (2004) realizou-se uma análise sobre a importância do momento central de ordem três no problema de otimização do portfólio de ativos de risco. Comparando esse modelo com o proposto por Markowitz (1952), procurou-se estabelecer quando um modelo é mais eficiente que o outro, sempre tendo como premissa as características dos mercados financeiros, preferências do investidor e conceitos gerais sobre economia.

Ficou evidente que o modelo de Markowitz (1952) é adequado em mercados financeiros que se comportam de modo similar à distribuição normal, de modo que o uso de uma função quadrática para retratar a utilidade é coerente. Entretanto, com o modelo de Athayde & Flôres Jr. (2004) foi observado que o momento central de ordem três pode auxiliar na escolha do conjunto de ativos que irá compor o portfólio do investidor, uma vez que, com uma assimetria positiva, o investidor está se precavendo de grandes perdas.

Foi realizada uma análise matemática da função utilidade, destacando como a preferência do investidor em exercer influência no comportamento dos termos da utilidade. Observou-se que o investidor tem preferência pelos momentos centrais de ordem ímpar e aversão aos momentos centrais de ordem par, devido aos momentos de ordem ímpar estarem ligados ao retorno sobre o portfólio, enquanto os momentos de ordem par são relativos à volatilidade presente no portfólio. Logo, o investidor possui aversão decrescente natural ao crescimento dos momentos de ordem par.

Pela análise dos termos da função utilidade, este trabalho contribui para um maior entendimento dos problemas de otimização propostos por Markowitz (1952) e Athayde & Flôres Jr. (2004). Revelou-se a consistência entre esses modelos e o princípio de maximização da utilidade, tendo como premissa os conceitos estabelecidos sobre a função utilidade e a preferência do investidor. Essa análise feita sobre a maximização da utilidade pode ser estendida para modelos que utilizem momentos centrais de altas ordens maiores que três.

Uma das contribuições de Markowitz foi mostrar que a otimização de um portfólio de ativos de risco não se trata apenas de aumentar o retorno sobre o portfólio, pois isso poderia ser feito de modo trivial, investindo a riqueza total no ativo com maior rentabilidade. Mas, ao fazer isso, o risco presente no portfólio aumentaria de forma absurda. Assim sendo, o investidor se colocaria em uma situação similar a de um jogo justo e, normalmente, o investidor é avesso a jogos justos por consequência da aversão ao risco. A partir disso percebe-se a necessidade da diversificação dos ativos financeiros de risco na carteira de investimentos. De modo a fundamentar a importância da diversificação num portfólio, realizou-se uma análise sobre a influência da mesma na maximização do portfólio. Foi observado, também, que com a diversificação obtém-se um portfólio ótimo. Em contrapartida, tornou-se possível reduzir a volatilidade presente no portfólio com o fracionamento da riqueza em diversos títulos, evitando, assim, uma situação de jogo justo.

Após discutir os modelos de seleção de portfólio, estabeleceu-se uma análise matemática sobre as aproximações da função utilidade considerando momentos centrais de ordem r. Ficou claro que uma aproximação da função utilidade que use momentos centrais de altas ordens indo até r é mais eficiente do que uma aproximação que adote momentos centrais de altas ordens menores que r. De modo que a solução encontrada para o problema de otimização do portfólio torne-se mais precisa. Com o estudo sobre as aproximações da utilidade, realizou-se uma crítica ao modelo de Markowitz, ressaltando um aspecto pouco considerado na discussão sobre a adequação de modelos de seleção de carteira. No caso de situações em que o risco presente no portfólio seja mínimo, isto é, seja próximo a zero, o modelo de Markowitz oferecerá resultados consistentes no problema de maximização da utilidade, mesmo que o mercado financeiro não se comporte semelhante a uma distribuição normal. Não obstante, ainda assim, se utilizarmos mais momentos em um caso que o risco tende a zero, teremos uma melhor aproximação da utilidade do que a proposta por Markowitz.

Nesse trabalho, a proposta era fazer um estudo sobre os fundamentos relativos à teoria de portfólio, com o intuito de trazer um maior entendimento sobre a mesma e revisar alguns conceitos. Foi alcançado, como resultado, a consolidação de uma coerência entre o problema de otimização proposto por Markowitz (1952) e Athayde & Flôres Jr. (2004) e o princípio de maximização da função utilidade. Enfatizamos a importância da diversificação, que é dita uma regra de ouro da economia (SAMUELSON, 1967). E, por fim, procuramos contribuir com a discussão em torno das aproximações da função utilidade esperada (SAMUELSON,1970).

REFERÊNCIAS

ASH, Robert B.. **Probability and Measury Theory.** Academic Press, 2000.

ATHAYDE, Gustavo M. de, & FLÔRES JR., Renato G.. **Finding a maximum skewness portfolio – a general solution to three-moments portfolio choice.** Journal of Economic Dynamics & Control, 2004.

BEAN, Michael A.. **Probability: The Science of Uncertainty with Applications to Investments, Insurance, and Engineering.** Brooks/Cole, 2001.

BORTOLOSSI, Humberto José. **Cálculo diferencial a várias variáveis:uma introdução à teoria de otimização.** Editora PUC-RIO, 2002.

CASTRO, Luciano I.de, & FARO, José Heleno. **Introdução à Teoria da Escolha.** IMPA, 2005.

COSTA NETO, Pedro Luiz de Oliveira. **Estatística.** Edgard Blucher, 1977.

FERNANDEZ, Pedro de Jesus. **Medida e Integração.** IMPA, 2002.

MAIBAUM, Gert. **Teoría de probabilidades y estadística matemática.** Pueblo y Educación, 1988.

MARKOWITZ, Harry. **Portfolio Selection.** The Journal of Finance, 1952.

NEUMANN, John von, & MORGENSTERN, Oskar. **Theory of Games and Economic Behavior.** Princeton University Press, 2004.

NIELSEN, Elena Kabatchenko. **Efficient Portfolio Selection in Mean-Variance-Skewness Space.** Ph.D. thesis, Arhus Universitet, 2008.

RUBINSTEIN, Mark. **Markowitz's "Portfolio Selection":A Fifty-Year Retrospective.** The Journal of Finance, 2002.

SAMUELSON, Paul A.. **General Proof that Diversification Pays.** The Journal of Finance and Quantitative Analysis, 1967.

SAMUELSON, Paul A.. **The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments.** The Review of Economic Studies, 1970.

SCDMITH, Klaus D.. **On the covariance of monotone functions of a random variable.** 2003. 1-3 f. Dresdner Schriften zur Versicherungsmathematik, 2003.

SCOTT, Robert C., & HORVATH, Philip A.. **On The Direction of Preference for Moments of Higher Order than the Variance.** The Journal of Finance, 1980.

SPIEGEL, Murray Ralph. Estatística: resumo da teoria, 875 problemas resolvidos, 619 problemas propostos. McGraw-Hill do Brasil, 1976.

APÊNDICE A – Conceitos e resultados de Teoria da Medida

Os conceitos abordados a seguir irão permitirão a definição dos conjuntos de Borel e a medida de Lebesgue-Stieltjes (ASH, 2000; FERNANDEZ, 2002). Com isso, será possível introduzir conceitos sobre probabilidade envolvendo os conjuntos de Borel e a função de Lebesgue.

Sejam $A_1, A_2, ...$ subconjuntos de um conjunto Ω. De acordo com a lei de De Morgan, temos que $(\bigcup_n A_n)^c = \bigcap_n A_n^c$ e $(\bigcap_n A_n)^c = \bigcup_n A_n^c$, onde A_n^c é usado para denotar o complementar de A_n em relação ao conjunto Ω.

Inicialmente, é preciso definir o conceito de σ-álgebra para, então, construir os conjuntos de Borel.

Definição A.0.1. (σ-álgebra) *Seja $\mathfrak{F}$ uma família de subconjuntos de um conjunto Ω. Diz-se que $\mathfrak{F}$ é uma σ-álgebra (sobre Ω) se $\Omega \in \mathfrak{F}$ e $\mathfrak{F}$ é fechada com relação ao complemento e a uniões enumeráveis, isto é,*

(a) $\Omega \in \mathfrak{F}$;

(b) Se $A \in \mathfrak{F}$, *então* $A^c \in \mathfrak{F}$;

(c) Se $A_1, A_2, ... \in \mathfrak{F}$, *então* $\cup_{i=1}^{\infty} A_i \in \mathfrak{F}$.

Note que a σ-álgebra $\mathfrak{F}$ é também fechada com relação a interseções enumeráveis. De fato, para $A_1, A_2, ... \in \mathfrak{F}$, temos $\cap_{i=1}^{\infty} A_i = (\cup_{i=1}^{\infty} A_i^c)^c \in \mathfrak{F}$.

Definição A.0.2. *Sejam Ω um conjunto e $\mathfrak{F}$ uma σ-álgebra de subconjuntos de Ω. Diz-se que o par $(\Omega, \mathfrak{F})$ é um espaço mensurável. Os elementos de $\mathfrak{F}$ são chamados de conjuntos mensuráveis.*

Definição A.0.3. *Sejam $(\Omega, \mathfrak{F})$ um espaço mensurável e μ uma função real estendida definida em $\mathfrak{F}$. Diz-se que μ é uma medida sobre*[7] $(\Omega, \mathfrak{F})$ *se μ satisfaz às condições:*

(a) $\mu(A) \in [0, \infty]$, *para* $A \in \mathfrak{F}$;

(b) $\mu(\emptyset) = 0$;

(c) Se $A_1, A_2, ... \in \mathfrak{F}$ *são conjuntos dois a dois disjuntos, então*

[7] Por vezes, será dito apenas que μ é uma medida sobre $\mathfrak{F}$.

$$\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

Se $\mu(\Omega) = 1$, diz-se que μ é uma medida de probabilidade.

Se $(\Omega, \mathfrak{F})$ é um espaço mensurável e μ é uma medida sobre $(\Omega, \mathfrak{F})$, diz-se que a tripla ordenada $(\Omega, \mathfrak{F}, \mu)$ é um espaço de medida. Quando μ é uma medida de probabilidade, $(\Omega, \mathfrak{F}, \mu)$ é chamado espaço de probabilidade.

Observação: A interseção de uma família qualquer de σ-álgebras é uma σ-álgebra. Portanto, dada uma família não-vazia $\mathfrak{F}$ de subconjuntos de um conjunto Ω, é possível definir a σ-álgebra gerada por $\mathfrak{F}$ como a menor σ-álgebra que contém $\mathfrak{F}$, e que, como é claro, coincide com a interseção de todas as σ-álgebras que contêm $\mathfrak{F}$. Quando Ω é um espaço topológico e $\mathfrak{F}$ é uma família dos subconjuntos abertos de Ω, a σ-álgebra gerada por $\mathfrak{F}$ é chamada de σ-álgebra de Borel do espaço topológico $(\Omega, \mathfrak{F})$ e é denotada por $\mathfrak{B}(\Omega)$. Os elementos dessa σ-álgebra são chamados de conjuntos de Borel.

Exemplo: Seja Ω o conjunto dos números reais estendidos, isto é, $\Omega = \overline{\mathbb{R}} = [-\infty, \infty]$. A σ-álgebra de Borel de $\overline{\mathbb{R}}$ é a gerada pelos intervalos abertos de $\overline{\mathbb{R}}$.

APÊNDICE B - Medidas de Lebesgue-Stieltjes e Funções de distribuição

Será construída uma classe de medidas nos conjuntos de Borel em $\mathbb{R}$, considerando funções F definidas de $\mathbb{R}$ em $\mathbb{R}$ que sejam não decrescentes e contínuas à direita. Isto é, se $a, b \in \mathbb{R}$ e $a < b$, então $F(a) \leq F(b)$ e, para todo $x_0 \in \mathbb{R}$, temos:

$$\lim_{x \to x_0^+} F(x) = F(x_0).$$

Definição A.0.4. *Uma medida de Lebesgue-Stieltjes em $\mathbb{R}$ é uma medida μ em $\mathfrak{B}(\mathbb{R})$, tal que $\mu(I) < \infty$ para cada intervalo I limitado. Uma função de distribuição em $\mathbb{R}$ é uma função $F: \mathbb{R} \to \mathbb{R}$ que é não-decrescente e contínua à direita.*

É possível mostrar que há uma correspondência um-para-um entre as medidas de Lebesgue-Stieltjes e funções de distribuição, se forem identificadas as distribuições que diferem por uma constante.

Teorema A.0.1. *Seja μ uma medida de Lebesgue-Stieltjes em $\mathbb{R}$. Seja $F: \mathbb{R} \to \mathbb{R}$ definida, a menos de uma constante aditiva, por $F(b) - F(a) = \mu(a, b]$. Então, F é uma função de distribuição.*

Teorema A.0.2. *Seja F uma função de distribuição em $\mathbb{R}$ e seja $\mu(a, b] = F(a, b], a, b \in \mathbb{R}, a < b$. Existe uma única extensão de μ para uma medida de Lebesgue-Stieltjes em $\mathbb{R}$.*

Serão consideradas agora funções F definidas de $\mathbb{R}^n$ em $\mathbb{R}$ e o operador diferença Δ. Seja $G: \mathbb{R}^n \to \mathbb{R}$, define-se $\Delta_{b_i a_i} G(x_1, \dots, x_n)$ por:

$$\Delta_{b_i a_i} G(x_1, \dots, x_n) = G(x_1, \dots, x_{i-i}, b_i, x_{i+1}, \dots, x_n) - G(x_1, \dots, x_{i-i}, a_i, x_{i+1}, \dots, x_n).$$

Definição A.0.5. *Se $a = (a_1, \dots, a_n)$ e $b = (b_1, \dots, b_n)$ pertencem a $\mathbb{R}^n$, o intervalo $(a, b]$ é definido como sendo o conjunto*

$$\{x = (x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x_i \leq b_i \; para \; todo \; i = 1, \dots, n\}.$$

Outros tipos de intervalos são definidos de maneira análoga.

Denota-se por $F(a,b]$ a seguinte composição:

$$F(a,b] = \Delta_{b_1 a_1} \cdots \Delta_{b_n a_n} F(x_1, \dots, x_n).$$

Entende-se que F é não decrescente se $F(a,b] \geq 0$ quando $a \leq b$ e que F é contínua à direita se F for contínua à direita em todas as componentes, simultaneamente. Isto é, para toda sequência (x^k) de vetores do $\mathbb{R}^n$ tais que $x^1 \geq x^2 \geq \cdots$ convergente a x, temos que $F(x^k)$ converge a $F(x)$.

Definição A.0.6. *Uma medida de Lebesgue-Stieltjes em $\mathbb{R}^n$ é uma medida μ em $\mathfrak{B}(\mathbb{R}^n)$, tal que $\mu(I) < \infty$ para cada intervalo I limitado. Uma função de distribuição em $\mathbb{R}^n$ é uma função $F: \mathbb{R}^n \to \mathbb{R}$ que é não-decrescente e contínua à direita.*

É possível mostrar que toda distribuição define uma medida de Lebesgue-Stieltjes.

Definição A.0.7. (Funções mensuráveis) *Sejam $(\Omega_1, \mathfrak{F}_1)$ e $(\Omega_2, \mathfrak{F}_2)$ espaços mensuráveis. Se $h: \Omega_1 \to \Omega_2$ é uma função, diz-se que h é mensurável com respeito às σ-*álgebras *$\mathfrak{F}_1$ e $\mathfrak{F}_2$ se $h^{-1}(A) \in \mathfrak{F}_1$ para todo $A \in \mathfrak{F}_2$. Se $\Omega_2 = \mathbb{R}^n$ e $\mathfrak{F}_2 = \mathfrak{B}(\mathbb{R}^n)$, diz-se que h é Borel mensurável*[8]*.*

---

[8] Quando $\boldsymbol{\Omega_1}$ é um conjunto de Borel de $\mathbb{R}^k$ (ou coincide com $\overline{\mathbb{R}^k}$), assume-se que $\boldsymbol{\mathfrak{F}_1} = \boldsymbol{\mathfrak{B}}(\mathbb{R}^k)$.