Carteiras de Investimentos e o Teorema do Melhor Fundo

Nesta palestra, desenvolvemos o famoso modelo de Markowitz para diversificação de carteiras de investimentos, introduzindo a análise de risco–retorno. Calculamos a fronteira eficiente através do uso de multiplicadores de Lagrange, provamos geometricamente (com um pouquinho de álgebra linear) o teorema dos dois fundos e, por fim, mostramos o teorema do melhor fundo, calculando uma carteira de investimentos ótima.

Na década de 1950, Harry Markowitz publicou seus estudos acerca de carteiras de investimentos, mostrando que a diversificação de investimentos permite reduzir substancialmente a variância inerente a uma carteira de ativos de risco. Seu trabalho pioneiro foi o marco-zero da moderna teoria de carteiras e serviu de base para o desenvolvimento de diversos outros modelos amplamente utilizados em Finanças, tais como o *Capital Asset Pricing Model* e a *Arbitrage Pricing Theory*.

A essência do modelo de Markowitz é aplicação ao retorno de ativos de risco de uma análise de média–variância. Dados n ativos de risco A_1, A_2, ... e A_n, podem-se calcular (utilizando de métodos estatísticos) os seus respectivos retornos esperados $\bar{r}_1$, $\bar{r}_2$, ... e $\bar{r}_n$, bem como seus respectivos desvios-padrões σ_1, σ_2, ... e σ_n. Isso permite representar cada os ativos em um diagrama de risco–retorno. Idealmente, um investidor deveria buscar ativos com o máximo de retorno e o mínimo de risco. Entretanto, geralmente os investimentos mais arriscados são justamente aqueles que oferecem o maior retorno (e vice-versa).

Markowitz percebeu que investindo em dois ativos distintos é possível obter uma carteira de investimento cujo retorno esperado é dado por uma média ponderada dos dois retornos esperados, mas com uma variância (e, portanto, com um desvio-padrão) menor do que a menor das duas variâncias. De fato, Markowitz percebeu que as covariâncias entre os retornos dos ativos desempenham um papel fundamental na diversificação de investimentos.

Para facilitar o entendimento do modelo de Markowitz, vamos usar a linguagem da álgebra linear. Definimos inicialmente o vetor de retornos esperados

$$\vec{R} = (\bar{r}_1, \bar{r}_2, ..., \bar{r}_n)$$

e a matriz de covariâncias

$$\mathbf{Q} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}$$

onde $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ e ρ_{ij} é o coeficiente de correlação entre os ativos A_i e A_j.

Uma carteira de invesstimentos será representada por um vetor

$$\vec{W} = (w_1, w_2, ..., w_n)$$

onde cada w_i representa o percentual do capital a ser investido no ativo A_i (valores negativos de w_i correspondem uma venda a descoberto do ativo A_i). Note que, por definção

$$w_1 + w_2 + ... + w_n = 1$$

O retorno esperado da carteira $\vec{W}$ será dado por um produto escalar:

$$\mathrm{Ret}(\vec{W}) = \vec{W} \cdot \vec{R} = \sum_{k=1}^{n} w_k \bar{r}_k$$

Por sua vez, a variância da carteira $\vec{W}$ será dado por:

$$\mathrm{Var}(\vec{W}) = \vec{W} \cdot (\mathbf{Q}\vec{W}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

Finalmente, o desvio-padrão será dado por:

$$\sigma(\vec{W}) = \sqrt{\mathrm{Var}(\vec{W})} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}$$

O conjunto factível $\mathcal{F}$ é o conjunto dos pontos do plano de risco–retorno que correspondem a carteiras formadas pelos ativos A_1, A_2, ... e A_n. Em símbolos,

$$\mathcal{F} = \Big\{ \Big(\sigma(\vec{W}), \mathrm{Ret}(\vec{W}) \Big) \in \mathbb{R}^2 : \vec{W} \in \mathbb{R}^n \;\text{ e }\; w_1 + w_2 + ... + w_n = 1 \Big\}$$

Em geral, a geometria do conjunto factível pode ser difícil de determinar. Todavia, estamos interessados apenas na fronteira esquerda do conjunto factível, conhecida como conjunto de variância mínima $\mathcal{V}$. Isso porque, dado um determinado nível de retorno esperado, desejamos investir na carteira de menor risco (i.e., a carteira mais à esquerda). Em símbolos,

$$\mathcal{V} = \big\{ (\sigma, \bar{r}) \in \mathbb{R}^2 : \sigma \leq \tilde{\sigma}, \; \forall \; (\tilde{\sigma}, \bar{r}) \in \mathcal{F} \big\}$$

Enquanto $\mathcal{F}$ é uma região do plano difícil de descrever, a geometria de $\mathcal{V}$ é extremamente simples: trata-se geralmente de uma curva clássica, a hipérbole.

Lema 1 — O conjunto de variância mínima $\mathcal{V}$ é um ramo de uma hipérbole (ou, no caso degenerado, um par de semi-retas).

Demonstração: Fixado um valor de retorno esperado $\bar{r}$, queremos encontrar o mínimo da função

$$\mathrm{Var}(\vec{W}) = \mathrm{Var}(w_1, w_2, ..., w_n) = \vec{W} \cdot \mathbf{Q}\vec{W} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

sujeita às restrições

$$\mathrm{Ret}(\vec{W}) = \mathrm{Ret}(w_1, w_2, ..., w_n) = \vec{W} \cdot \vec{R} = \sum_{k=1}^{n} w_k \bar{r}_k = \bar{r}$$

e

$$f(w_1, w_2, ..., w_n) = \vec{W} \cdot \vec{1} = w_1 + w_2 + ... + w_n = 1$$

onde $\vec{1} = (1, 1, ..., 1)$. Pelo método dos multiplicadores de Lagrange, basta resolvermos a equação

$$\nabla\text{Var}(w_1, w_2, ..., w_n) = \lambda\nabla\text{Ret}(w_1, w_2, ..., w_n) + \mu\nabla f(w_1, w_2, ..., w_n)$$

sujeita às duas restrições acima. Calculando

$$\nabla\text{Var}(w_1, w_2, ..., w_n) = \left(2\sum_{k=1}^{n} w_k\sigma_{1k}, 2\sum_{k=1}^{n} w_k\sigma_{2k}, ..., 2\sum_{k=1}^{n} w_k\sigma_{nk}\right)$$

e notando que $\nabla\text{Ret}(w_1, w_2, ..., w_n) = \vec{R}$ e $\nabla f(w_1, w_2, ..., w_n) = \vec{1}$, temos que o ponto procurado corresponde à solução do sistema linear

$$\left\{\begin{array}{ccccccccccccc}
2\sigma_{11}w_1 & + & 2\sigma_{12}w_2 & + & ... & + & 2\sigma_{1n}w_n & - & \bar{r}_1\lambda & - & \mu & = & 0 \\
2\sigma_{21}w_1 & + & 2\sigma_{12}w_2 & + & ... & + & 2\sigma_{2n}w_n & - & \bar{r}_2\lambda & - & \mu & = & 0 \\
\vdots & & \vdots & & \vdots & & & & & & & \vdots & \\
2\sigma_{n1}w_1 & + & 2\sigma_{n2}w_2 & + & ... & + & 2\sigma_{nn}w_n & - & \bar{r}_n\lambda & - & \mu & = & 0 \\
\bar{r}_1w_1 & + & \bar{r}_2w_2 & + & ... & + & \bar{r}_nw_n & & & & & = & \bar{r} \\
w_1 & + & w_2 & + & ... & + & w_n & & & & & = & 1
\end{array}\right.$$

Note que, ao resolver o sistema, a solução será da forma

$$V(\bar{r}) = (a_1 + b_1\bar{r}, a_2 + b_2\bar{r}, ..., a_n + b_n\bar{r})$$

onde a_1, a_2, ..., a_n, b_1, b_2, ... e b_n são constantes reais que não dependem de $\bar{r}$. Assim, a variância mínima correspondente a $\bar{r}$ será dada por:

$$V(\bar{r}) \cdot \mathbf{Q}V(\bar{r}) = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_i + b_i\bar{r})(a_j + b_j\bar{r})\sigma_{ij}$$

Rearranjando os termos, temos:

$$V(\bar{r})\cdot\mathbf{Q}V(\bar{r}) = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} b_ib_j\sigma_{ij}\right)\bar{r}^2 + \left(\sum_{i=1}^{n}\sum_{j=1}^{n}(a_ib_j + a_jb_i)\sigma_{ij}\right)\bar{r} + \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_ia_j\sigma_{ij}\right)$$

Portanto, o desvio padrão dacarteira de variância mínima correspondente a $\bar{r}$ será da forma

$$\sigma(\bar{r}) = \sqrt{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j \sigma_{ij}\right)\bar{r}^2 + \left(\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j + a_j b_i)\sigma_{ij}\right)\bar{r} + \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \sigma_{ij}\right)}$$

Uma vez que a variância nunca é negativa, o radicando será sempre não negativo. Conseqüentemente, o gráfico de $\sigma(\bar{r})$ será um ramo de uma hipérbole ou, no caso de o radicando se anular para algum $\bar{r}$, um par de semi-retas. $\square$

Observamos que o caso degenerado simplesmente não ocorre na prática, uma vez que a existência de uma carteira sem risco composta por ativos de risco é altamente improvável. De fato, para que isso fosse possível, deveríamos ter ativos de risco perfeitamente correlacionados ou negativamente perfeitamente correlacionados, o que é um evento de probabilidade extremamente baixa por motivos econômicos.

O método de multiplicadores de Lagrange, usado no Lema 1 para calcular a fronteira eficiente, raramentente é empregado para esse fim, em virtude do teorema dos dois fundos.

De fato, da demonstração do Lema 1 decorre que o conjunto de variância mínima $\mathcal{V}$ é usualmente o ramo direito de uma hipérbole com eixo focal hoizontal. O vértice dessa hipérbole corresponde à carteira de variância mínima, i.e., à carteira de investimentos de menor risco que pode ser formada combinando os ativos de risco. A metade superior de $\mathcal{V}$ é chamada de fronteira eficiente $\mathcal{E}$. A fronteira eficiente $\mathcal{E}$ é o conjunto formado pelas carteiras que tem o máximo retorno esperado para cada nível de risco escolhido.

Teorema dos Dois Fundos — Dadas duas carteiras distintas sobre o conjunto de variância mínima $\mathcal{V}$, qualquer carteia sobre esse conjunto pode ser obtida investindo em uma combinação somente dessas duas carteiras.

Demonstração: Dadas duas carteiras distintas F_1 e F_2 sobre o conjunto de variância mínima $\mathcal{V}$, aplicando o Lema 1 para apenas dois ativos (F_1 e F_2), vemos que as combinações dessas duas carteiras descrevem uma hipérbole $\mathcal{H}$ com eixo focal hoizontal. Vamos provar que $\mathcal{H} = \mathcal{V}$. De fato, $\mathcal{H}$ e $\mathcal{V}$ são duas hipérboles com eixo focal hoizontal que possuem dois pontos em comum, F_1 e F_2. Note que, se $\mathcal{H}$ interceptasse $\mathcal{V}$ transversalmente, teríamos pontos de $\mathcal{H}$ à esquerda de $\mathcal{V}$, o que contradiz a definição de $\mathcal{V}$ como conjunto de variância mínima. Potanto, $\mathcal{H}$ e $\mathcal{V}$ são tangentes em F_1 e F_2. Pela demonstração do Lema 1,

$$\mathcal{H} : \sigma = \sqrt{h_1\bar{r}^2 + h_2\bar{r} + h_3}$$

e

$$\mathcal{V} : \sigma = \sqrt{v_1\bar{r}^2 + v_2\bar{r} + v_3}$$

Sendo $F_1 = (\bar{r}_1, \sigma_1)$ e $F_2 = (\bar{r}_2, \sigma_2)$, temos que

$$\begin{cases} h_1\bar{r}_1^2 + h_2\bar{r}_1 + h_3 = v_1\bar{r}_1^2 + v_2\bar{r}_1 + v_3 \\ 2h_1\bar{r}_1 + h_2 = 2v_1\bar{r}_1 + v_2 \\ h_1\bar{r}_2^2 + h_2\bar{r}_2 + h_3 = v_1\bar{r}_2^2 + v_2\bar{r}_2 + v_3 \\ 2h_1\bar{r}_2 + h_2 = 2v_1\bar{r}_2 + v_2 \end{cases}$$

Da segunda e da quarta equações temos que, se $v_1 \neq h_1$:

$$\bar{r}_1 = -\frac{v_2 - h_2}{2(v_1 - h_1)} = \bar{r}_2$$

o que é absurdo, visto que F_1 e F_2 são distintos. Logo, $v_1 = h_1$. Da primeira e da terceira equações temos que:

$$(v_1 - h_1)\bar{r}^2 + (v_2 - h_2)\bar{r} + (v_3 - h_3) = (v_1 - h_1)(\bar{r} - \bar{r}_1)(\bar{r} - \bar{r}_2)$$

e, portanto, $v_2 = h_2$ e $v_3 = h_3$, de modo que $\mathcal{H} = \mathcal{V}$. $\square$

Em vista do teorema dos dois fundos, basta encontrarmos duas carteiras distintas quaisquer sobre o conjunto de variância mínima, ou seja, duas soluções quaisquer do sistema dado no Lema 1. Uma solução particularmente fácil de calcular corresponde ao caso $\lambda = 0$ e $\mu = 2$, quando as n primeiras equações do sistema reduzem-se a:

$$\mathbf{Q}\vec{U} = \vec{1}$$

A carteira $\vec{W}_1$, correspondente ao primeiro fundo F_1, é então obtida normalizando a solução U do sistema acima:

$$\vec{W}_1 = \frac{\vec{U}}{\vec{U} \cdot \vec{1}}$$

Outra solução simples de calcular corresponde ao caso $\lambda = 2$ e $\mu = 0$, quando as n primeiras equações do sistema reduzem-se a:

$$\mathbf{Q}\vec{V} = \vec{R}$$

A carteira $\vec{W}_2$, correspondente ao segundo fundo F_2, é então obtida normalizando a solução V do sistema acima:

$$\vec{W}_2 = \frac{\vec{V}}{\vec{V} \cdot \vec{1}}$$

Qualquer fundo sobre o conjunto de variância mínima pode assim ser escrito na forma

$$\vec{V}_\alpha = \alpha\vec{W}_1 + (1-\alpha)\vec{W}_2$$

onde $\alpha \in \mathbb{R}$. O retorno esperado dessa carteira será

$$\mathrm{Ret}(\vec{V}_\alpha) = \alpha\mathrm{Ret}(\vec{W}_1) + (1-\alpha)\mathrm{Ret}(\vec{W}_2)$$

Essa fórmula permite-nos encontrar o valor de α correspondente a um dado nível de retorno esperado $\bar{r}$. De fato,

$$\alpha = \frac{\bar{r} - \mathrm{Ret}(\vec{W}_2)}{\mathrm{Ret}(\vec{W}_1) - \mathrm{Ret}(\vec{W}_2)}$$

Desse modo, encontramos a carteira de menor variância para cada nível de retorno esperado escolhido.

Consideremos agora um ativo sem risco T de retorno fixo r_f. Tal ativo é geralmente um título do Tesouro Nacional. Ora, note que a covariância desse retorno com o retorno de qualquer ativo de risco é nula. Uma combinação desse ativo sem risco com uma carteira $\vec{W}$ de ativos de risco será da forma

$$V_t = t\vec{W} + (1-t)T$$

com retorno esperado

$$\mathrm{Ret}(V_t) = t\mathrm{Ret}(\vec{W}) + (1-t)r_f$$

e variância

$$\mathrm{Var}(V_t) = t^2\mathrm{Var}(\vec{W})$$

Assim, correspondente a $t \geq 0$ no diagrama de risco–retorno, temos a semi-reta que tem origem no ponto $(0, r_f)$ correspondente a T e passa pelo ponto $(\sigma(\vec{W}), \mathrm{Ret}(\vec{W}))$. O coeficiente angular dessa semi-reta é

$$\tan\theta = \frac{\mathrm{Ret}(\vec{W}) - r_f}{\sigma(\vec{W})}$$

À medida que a carteira $\vec{W}$ de ativos de risco percorre o conjunto de variância mínima, as combinações de $\vec{W}$ com o ativo sem risco T geram um novo conjunto factível e, conseqüentemente, uma nova fronteira eficiente. Essa nova fronteira eficiente será a semi-reta com origem no ativo sem risco T e que tangencia superiormente o conjunto de variância mínima $\mathcal{V}$.

Teorema do Melhor Fundo — Esixte uma única carteira $\vec{M}$ de ativos de risco, tal que qualquer carteira sobre a fronteira eficiente pode ser obtida investindo em uma combinação somente dessa carteira com o ativo sem risco T.

Demonstração: A carteira procurada correspondente ao ponto de máximo do coeficiente angular

$$\tan\theta = \frac{\mathrm{Ret}(\vec{W}) - r_f}{\sigma(\vec{W})}$$

A existência desse ponto de máximo é garantida, desde que o retorno do ativo sem risco seja menor do que o retorno da cateira de variância mínima. (Em virtude do equilíbrio econômico, essa condição é sempre verificada na prática.) Substituindo as expressões de $\mathrm{Ret}(\vec{W})$ e $\sigma(\vec{W})$, temos

$$\tan\theta = \frac{\displaystyle\sum_{k=1}^{n} w_k(\bar{r}_k - r_f)}{\sqrt{\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}}}$$

No ponto de máximo, as derivadas parciais devem-se anular. Assim, para cada $k = 1, 2, ..., n$, devemos ter

$$\frac{\partial}{\partial w_k}\tan\theta = \frac{(\bar{r}_k - r_f)\mathrm{Var}(\vec{W}) - \Big(\mathrm{Ret}(\vec{W}) - r_f\Big)\displaystyle\sum_{i=1}^{n} w_i\sigma_{ik}}{\Big[\sigma(\vec{W})\Big]^3} = 0$$

ou seja

$$\frac{\mathrm{Ret}(\vec{W}) - r_f}{\mathrm{Var}(\vec{W})}\sum_{i=1}^{n} w_i\sigma_{ik} = \bar{r}_k - r_f$$

Fazendo a mudança de variáveis

$$\tilde{w}_k = \frac{\mathrm{Ret}(\vec{W}) - r_f}{\mathrm{Var}(\vec{W})} w_k$$

obtemos, para cada $k = 1, 2, ..., n$,

$$\sum_{i=1}^{n} \tilde{w}_i \sigma_{ik} = \bar{r}_k - r_f$$

Desse modo, o ponto de máximo será a solução nomalizada do sistema

$$\mathbf{Q}\vec{W} = \vec{R} - r_f \vec{1}$$

De acordo com a discussão feita após a demonstração do Teorema dos Dois Fundos, a solução desse sistema é

$$\vec{W} = \vec{V} - r_f \vec{U}$$

e o melhor fundo será dado por

$$\vec{M} = \frac{\vec{V} - r_f \vec{U}}{\vec{V} \cdot \vec{1} - r_f \left(\vec{U} \cdot \vec{1} \right)}$$

□

Observe que os fundos obtidos após o Teorema dos Dois Fundos podem ser obtidos como casos particulares do melhor fundo obtido acima, tomando $r_f = 0$ e fazendo $r_f \to \infty$. Com efeito, o fundo F_1 corresponde à carteira de vaiância mínima, enquanto o fundo F_2 é o melhor fundo para uma taxa de juros sem risco nula.

Em vista do Teorema do Melhor Fundo, tudo que um investidor precisa fazer é repartir seu capital entre o ativo sem risco e o melhor fundo, de acordo com o nível de risco ao qual ele deseja se expor. Por sua vez, tudo que um bom banco de investimento precisa oferecer a seus clientes é um bom título de renda fixa (CDB) e um bom fundo de investimento (FIF).