Diversificação de Portfólios e Otimização

Preliminares: Conforme definimos no capítulo xxx o portfólio é dado pelo vetor coluna:

$$\vec{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix}$$

De um investidor com patrimônio [wealth] $w_o = n_1 S_1 + n_2 S_2 + \ldots + n_A S_A$, onde n_j é o número de ações na jésima ação cujo preço é dado por S_j, é dado pelo vetor de frações do investidor em cada um dos ativos $x_j = \frac{n_j S_j}{W_o}$. Obviamente que o vetor portfólio obedece à restrição $\sum_i x_i = 1$. O x pode ser positivo, “long position”, com o significado de que o investidor comprou a ação, ou negativo, “short position”, com o significado de que o investidor fica devendo a ação.

O retorno de uma ação é definido como $R_t = \frac{p_t - p_{t-1}}{p_{t-1}}$ e o log-retorno como $r_t = Ln(1 + R_t) = Ln\left(\frac{p_t}{p_{t-1}}\right)$. A covariância e correlação entre as ações é obtida das séries temporais do log-retorno.

Vamos considerar agora o rendimento e a variância de um portfólio $\vec{X}$. O rendimento do mesmo é dado por $r = \sum_i x_i r_i$ cuja esperança vale $\bar{r} = \sum_i x_i \bar{r}_i$, onde $\bar{Z} = E[Z]$. Dessa forma $E\left[(r - \bar{r})^2\right] = \sum_i \sum_j x_i E\left[(r_i - \bar{r}_i)(r_j - \bar{r}_j)\right] x_j = \vec{X}\, V\, \vec{X}^T$ onde $\vec{X}^T$ é o vetor coluna de transposto de $\vec{X}$ e $V_{ij} = E\left[(r_i - \bar{r}_i)(r_j - \bar{r}_j)\right]$ é a matriz de variância-covariância, denominada daqui em diante por matriz de covariância. Também sabemos do capítulo xxx que o coeficiente de correlação é definido como $r_{ij} = \frac{\operatorname{cov}(r_i, r_j)}{\sigma_i \sigma_j}$, logo $\operatorname{cov}(r_i, r_j) = r_{ij}\sigma_i\sigma_j$ e sabemos que $-1 \le r_{ij} \le +1$.

Vamos tomar um exemplo do poder da diversificação com apenas dois ativos arriscados. O ativo 1 apresenta $\overline{r}_1$ e desvio padrão σ_1 enquanto o atrivo 2 apresenta $\overline{r}_2$ e desvio padrão σ_2 e a covariância entre os dois $V_{12} = \sigma_{12}$. Em lugar de investir tudo em qulquer um dos dois ativos vamos investir a fração x_1 no ativo 1 e a fração x_2 no ativo 2, de forma que $x_1 + x_2 = 1$. Nesse caso o retorno do portfólio será dado por:

$$r = x_1 r_1 + x_2 r_2$$

O retorno médio $\overline{r} = E[r] = x_1 E[r_1] + x_2 E[r_2] = x_1 \overline{r}_1 + x_2 \overline{r}_2$. Logo:

$$r - \overline{r} = x_1 (r_1 - \overline{r}_1) + x_2 (r_2 - \overline{r}_2)$$

Dessa forma:

$$(r - \overline{r})^2 = x_1^2 (r_1 - \overline{r}_1)^2 + 2x_1 x_2 (r_1 - \overline{r}_1)(r_2 - \overline{r}_2) + x_2^2 (r_2 - \overline{r}_2)^2$$

Nesse caso a variância do portfólio será dada por:

$$\sigma_r^2 = E\left[(r - \overline{r})^2\right] = x_1^2 E\left[(r_1 - \overline{r}_1)^2\right] + 2x_1 x_2 E\left[(r_1 - \overline{r}_1)(r_2 - \overline{r}_2)\right] + x_2^2 E\left[(r_2 - \overline{r}_2)^2\right]$$

Ou seja:

$$\sigma_r^2 = x_1^2 \sigma_1^2 + 2x_1 x_2 \operatorname{cov}(r_1, r_2) + x_2^2 \sigma_2^2$$

Que pode ser escrita em termos do coeficiente de correlação como:

$$\sigma_r^2 = x_1^2 \sigma_1^2 + 2x_1 x_2 r_{12} \sigma_1 \sigma_2 + x_2^2 \sigma_2^2$$

Vamos considerar dois casos especiais.

1. Suponha um caso especial no qual $r_{12} = +1$, então $\sigma_r^2 = x_1^2 \sigma_1^2 + 2x_1 x_2 \sigma_1 \sigma_2 + x_2^2 \sigma_2^2 = (x_1 \sigma_1 + x_2 \sigma_2)^2$ então $\sigma_r = x_1 \sigma_1 + x_2 \sigma_2$.

Como $\overline{r} = x_1\overline{r}_1 + x_2\overline{r}_2$ também, variando as frações pode-se deslocar ao longo da reta que une os pontos $(\sigma_1, \overline{r}_1)$ ao ponto $(\sigma_2, \overline{r}_2)$.

2. Suponha agora outra situação especial de anticorrelação perfeita na qual $r_{12} = -1$. Nesse caso $\sigma_r^2 = x_1^2\sigma_1^2 - 2x_1x_2\sigma_1\sigma_2 + x_2^2\sigma_2^2 = (x_1\sigma_1 - x_2\sigma_2)^2$ então $\sigma_r = |x_1\sigma_1 - x_2\sigma_2|$. Note que nesse caso é possível eliminar completamente o risco escolhendo as frações de forma que $x_1\sigma_1 - x_2\sigma_2 = 0$, $x_2 = \frac{\sigma_1}{\sigma_2}x_1$, logo $x_1 + x_2 = x_1 + \frac{\sigma_1}{\sigma_2}x_1 = \frac{\sigma_1+\sigma_2}{\sigma_2}x_1 = 1$, o que significa que $x_1 = \frac{\sigma_2}{\sigma_1+\sigma_2}$ e $x_2 = \frac{\sigma_1}{\sigma_1+\sigma_2}$. Nesse ponto $\overline{r} = x_1\overline{r}_1 + x_2\overline{r}_2 = \frac{\sigma_2\overline{r}_1 + \sigma_1\overline{r}_2}{\sigma_1+\sigma_2}$ é uma média ponderada dos dois retornos com os riscos como pesos. Se $x_1\sigma_1 - x_2\sigma_2 < 0$ então:

$$\sigma_r = x_2\sigma_2 - x_1\sigma_1 = (1-x_1)\sigma_2 - x_1\sigma_1 = \sigma_2 - x_1(\sigma_1+\sigma_2)$$

Logo $x_1 = \frac{\sigma_2 - \sigma_r}{\sigma_1+\sigma_2}$ e a curva $\overline{r}\ vs\ \sigma$ pode ser obtida de $\overline{r} = x_1\overline{r}_1 + (1-x_1)\overline{r}_2 = \overline{r}_2 + x_1(\overline{r}_1 - \overline{r}_2)$ que resulta na reta:

$$\overline{r} = \frac{\sigma_1\overline{r}_2 + \sigma_2\overline{r}}{\sigma_1+\sigma_2} - \frac{\overline{r}_1 - \overline{r}_2}{\sigma_1+\sigma_2}\sigma_r$$

É uma reta negativamente inclinida com valor no intercepto de $\overline{r}_{\text{int}} = \frac{\sigma_1\overline{r}_2 + \sigma_2\overline{r}}{\sigma_1+\sigma_2}$ extamente o $\overline{r}$ no risco zero. Por outro lado se $x_1\sigma_1 - x_2\sigma_2 > 0$ então $\sigma_r = x_1\sigma_1 - x_2\sigma_2 = x_1\sigma_1 - (1-x_1)\sigma_2 = x_1(\sigma_1+\sigma_2) - \sigma_2$ logo $x_1 = \frac{\sigma_2+\sigma_r}{\sigma_1+\sigma_2}$ e a reta muda para $\overline{r} = \frac{\sigma_1\overline{r}_2 + \sigma_2\overline{r}}{\sigma_1+\sigma_2} + \frac{\overline{r}_1 - \overline{r}_2}{\sigma_1+\sigma_2}\sigma_r$, poistivamente inclinada passando no ponto $(0,\bar{r}_{\text{int}})$. Esse caso especial mostra então que foi possível elimininar completamente o risco e obter um retorno entre os dois retornos individuais.

3. Finalmente vamos supor a situação de dois ativos descorrelacionados, na qual $r_{12}=0$. Nesse caso $\sigma_r^2 = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 = x_1^2\sigma_1^2 + (1-x_1)^2\sigma_2^2 = x_1^2\sigma_1^2 + \sigma_2^2 - 2x_1\sigma_2^2 + x_1^2\sigma_2^2$, ou seja, $\sigma_r^2 = x_1^2\left(\sigma_1^2+\sigma_2^2\right) - 2x_1\sigma_2^2 + \sigma_2^2$ agora portanto:

$$\sigma_r^2 = \left(\sigma_1^2+\sigma_2^2\right)\frac{\left(\bar{r}-\bar{r}_2\right)^2}{\left(\bar{r}_1-\bar{r}_2\right)^2} - 2\frac{\left(\bar{r}-\bar{r}_2\right)}{\left(\bar{r}_1-\bar{r}_2\right)}\sigma_2^2 + \sigma_2^2$$

$$\frac{\sigma_r^2}{\sigma_1^2+\sigma_2^2} = \left(\frac{\bar{r}-\bar{r}_2}{\bar{r}_1-\bar{r}_2}\right)^2 - 2\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)\left(\frac{\bar{r}-\bar{r}_2}{\bar{r}_1-\bar{r}_2}\right) + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$$

$$\frac{\sigma_r^2}{\sigma_1^2+\sigma_2^2} = \left(\frac{\bar{r}-\bar{r}_2}{\bar{r}_1-\bar{r}_2}\right)^2 - 2\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)\left(\frac{\bar{r}-\bar{r}_2}{\bar{r}_1-\bar{r}_2}\right) + \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2 - \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2 + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$$

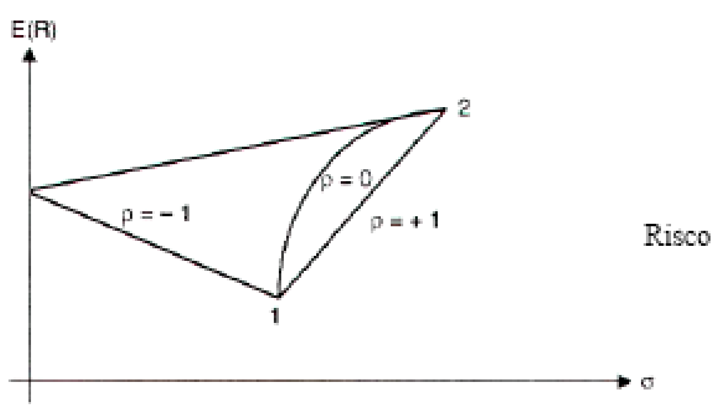
$$\frac{\sigma_r^2}{\sigma_1^2+\sigma_2^2} = \left(\frac{\bar{r}-\bar{r}_2}{\bar{r}_1-\bar{r}_2} - \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2 + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\left(1-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)$$

$$\frac{\sigma_r^2}{\sigma_1^2+\sigma_2^2} = \left(\frac{\bar{r} - \dfrac{\sigma_2^2\bar{r}_1+\sigma_1^2\bar{r}_2}{\sigma_1^2+\sigma_2^2}}{\left(\bar{r}_1-\bar{r}_2\right)}\right)^2 + \frac{\sigma_1^2\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2\right)^2}$$

$$\frac{\sigma_r^2}{\left(\dfrac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2} - \frac{\left(\bar{r} - \dfrac{\sigma_2^2\bar{r}_1+\sigma_1^2\bar{r}_2}{\sigma_1^2+\sigma_2^2}\right)^2}{\left(\dfrac{\sigma_1\sigma_2\left(\bar{r}_1-\bar{r}_2\right)}{\sigma_1^2+\sigma_2^2}\right)^2} = 1$$

Que é um hipérbole com vértice em $\sigma_r = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$ e $\bar{r}_{\text{int}} = \frac{\sigma_2^2\bar{r}_1 + \sigma_1^2\bar{r}_2}{\sigma_1^2+\sigma_2^2}$.

Figura xxx ilustra os três casos. Os casos correlação/anticorrelação perfeitas $r_{12} = \pm 1$ são, na prática, impossíveis. Mas mesmo assim pode-se perceber a vantagem de diversificar o portfólio para diminuição dos riscos. Mesmo no caso descorrelacionado é possível diminuir o risco para $\sigma_r = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$ obtendo o retorno $\bar{r}_{\text{int}} = \frac{\sigma_2^2\bar{r}_1 + \sigma_1^2\bar{r}_2}{\sigma_1^2+\sigma_2^2}$. Note que esse problema pode ser colocado como um problema de encontrar o x para o qual a variância $\sigma_r^2 = x^2\sigma_1^2 + (1-x)^2\sigma_2^2$ é mínima. Nesse caso $\frac{d}{dx}\sigma_r^2 = 2x\left(\sigma_1^2+\sigma_2^2\right) - 2\sigma_2^2 = 0$ e $x = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ com $1-x = \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}$ que era o mesmo que que anulava o risco no caso $r_{12} = -1$. Só que agora $\sigma_r^2 = \frac{\sigma_1^2\sigma_2^4}{\left(\sigma_1^2+\sigma_2^2\right)^2} + \frac{\sigma_1^4\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2\right)^2} = \sigma_1^2\sigma_2^2\left[\frac{\sigma_1^2+\sigma_2^2}{\left(\sigma_1^2+\sigma_2^2\right)^2}\right] = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$ e $\sigma_r = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$.



Propriedades da Matriz de Covariância $V_{ij} = E\left[(r_i - \overline{r}_i)(r_j - \overline{r}_j)\right]$

1. A matriz V_{ij} é simétrica:

$$V_{ij} = E\left[(r_i - \overline{r}_i)(r_j - \overline{r}_j)\right] = E\left[(r_j - \overline{r}_j)(r_i - \overline{r}_i)\right] = V_{ji}$$

2. A matriz V_{ij} é definida positiva.

Demonstração*:* Do capítulo de matrizes e álgebra linear sabemos que uma matriz P é definida positiva se $q = \vec{X}'M\vec{X} > 0$, para qualquer $\vec{X} \neq \vec{O}$, onde $\vec{X}'$ é o vetor linha transposto de $\vec{X}$. Note que q é um escalar pois $q_{1\times1} = \vec{X}'_{1\times n} M_{n\times n} \vec{X}_{n\times1}$. Essa condição pode ser escrita como $q = \sum_i \sum_j x_i M_{ij} x_j > 0$. Para mostrar que a matriz de variância-covariância é definida positiva considere a grandeza z definida por $z = \sum_i x_i (r_i - \overline{r}_i)$. Nesse caso $z^2 = \sum_i \sum_j x_i x_j (r_i - \overline{r}_i)(r_j - \overline{r}_j)$, e $E[z^2] = \sum_i \sum_j x_i E\left[(r_i - \overline{r}_i)(r_j - \overline{r}_j)\right] x_j$, ou seja, $E[z^2] = \sum_i \sum_j x_i V_{ij}\, x_j$. Porém, $z^2 > 0$ sempre, logo $E[z^2] > 0$, que implica que $\sum_i \sum_j x_i V_{ij}\, x_j > 0$ para $\forall \vec{X}$, ou seja, que V_{ij} é definida positiva.

3. Se V_{ij} = V_{ji} é simétrica então $V_{ij}^{-1} = V_{ji}^{-1}$também é simétrica.

Demonstração: Sabemos que $VV^{-1} = I$, aplicando a transposta de ambos os lados temos $(V^{-1})^T V^T = I^T$. Mas $I^T = I$ sempre e $V^T = V$ por hipótese, logo $(V^{-1})^T V^T = I$, o que significa que $(V^{-1})^T = V^{-1}$.

4. Se V é simétrica e definida positiva então V^{-1} também é simétrica e definida positiva.

Demonstração: Sabemos que $\vec{X}'V\vec{X} > 0$, pois V é definida positiva para $\forall\,\overrightarrow{X}$.

Simetria: Pela definição de matriz inversa $V^{-1}V = I$ e $V V^{-1} = I$. Aplicando a transposta na primeira identidade obtemos $\overline{V^{-1}V} = \overline{I}$ ou seja $\overline{V}\,\overline{V^{-1}} = I$. Mas $\overline{V} = V$ pois V é simétrica, então $V\overline{V^{-1}} = I$, o que significa que $\overline{V^{-1}} = V^{-1}$, ou seja, a inversa de uma matriz simétrica também é simétrica.

Definida positiva: Agora, $\vec{X}'V\vec{X} = \vec{X}'VV^{-1}\underbrace{VV^{-1}}_{I}V\vec{X} = \left(\vec{X}'V\right)V^{-1}\left(V\vec{X}\right) = \left(\overline{V\vec{X}}\right)V^{-1}\left(V\vec{X}\right)$. Mas

Se $V\vec{X} = \vec{Y}$ é um vetor arbitrário então, $\vec{Y}'V^{-1}\vec{Y} = \vec{X}'V\vec{X} > 0$, logo $\vec{Y}'V^{-1}\vec{Y} > 0$ para $\forall \vec{Y}$, o que significa que V^{-1} também é definida positiva.

Modelo de Markowitz [http://en.wikipedia.org/wiki/Harry_Markowitz]

Harry Markowitz [1927] ganhou o prêmio Nobel de Economia de 1990 pelo seu trabalho de tese de Doutorado na Universidade de Chicago em 1955, a qual fundou a área hoje denominada por Modern Portfolio Theory [MPT]. A Universidade de Chicago decidiu em 1942 dar um diploma de Bachelor a seus estudantes após dois anos de estudos sem uma formação específica. Como se trata de uma das melhores universidades do mundo, com 89 prêmios Nobel, seu exemplo se provou um sucesso o qual deveria ser copiado em todo o mundo. Antes de se decidir por economia Markowitz graduou-se nesse programa de dois anos mostrando interesse especial em filosofia e física. Milton Friedman, prêmio Nobel de economia da Universidade de Chicago, fazia parte da sua banca de defesa de tese e chegou a afirmar que seu trabalho, o qual rendeu prêmio Nobel, não era economia.

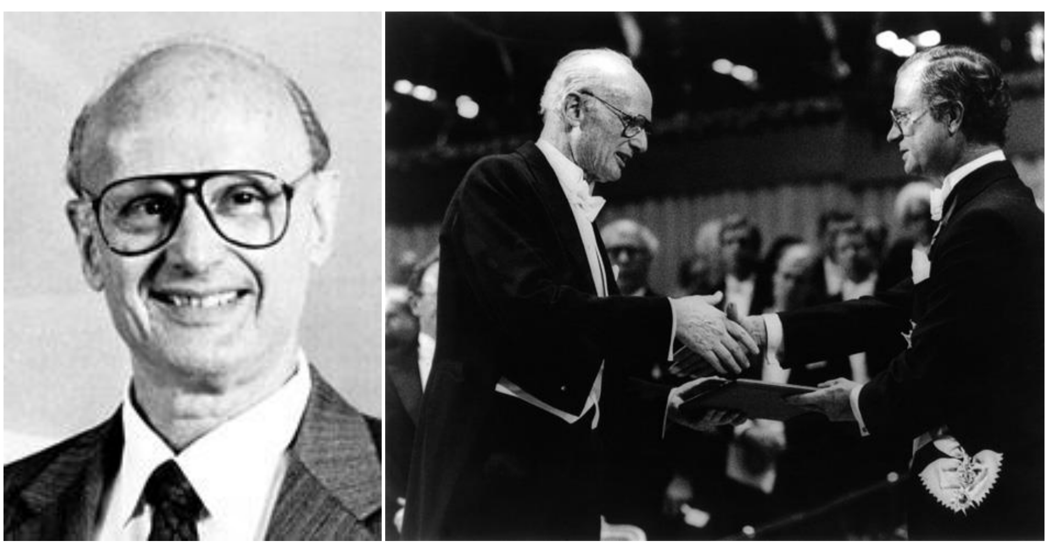


Figura xxx. Harry Markowitz. Recebendo o prêmio Nobel em Estocolmo.

As hipóteses do modelo de Markowitz são:

1. Investidor maximiza a utilidade esperada da riqueza ao final de um período
2. Investidor tem aversão ao risco
3. Escolha do investidor é baseada apenas em termos do retorno médio e da variância, utilizada como proxy do risco
4. Mercados financeiros são perfeitos – não existem custos de transação, os títulos são infinitamente divísiveis e os investidores são "price-takers", sem poder de interferir nos preços das ações.
5. Todos os ativos são arriscados e nenhum par de ativos apresenta coeficiente de correlação ± 1.
6. A hipótese mais importante do modelo de Markowitz recai sobre as restrições: além de exigir que $\sum_i x_i = 1$ ele requer que $x_i \geq 0$ para $\forall i$. Ou seja, não permite a short position.

A fronteira eficiente é definida como a curva retorno versus risco possível com o conjunto de ativos arriscados dado. O detalhe importante nesse modelo é que tanto a matriz de covariância quanto os retornos médios que importam para o investidor são os do próximo período, no futuro, em princípio não conhecidos. A hipótese mais forte do modelo é que o futuro será uma repetição do passado, então os retornos e a matriz de covariância históricos representarão bem o futuro.

O modelo de Markowitz é expresso, então, como um problema de otimização dado por:

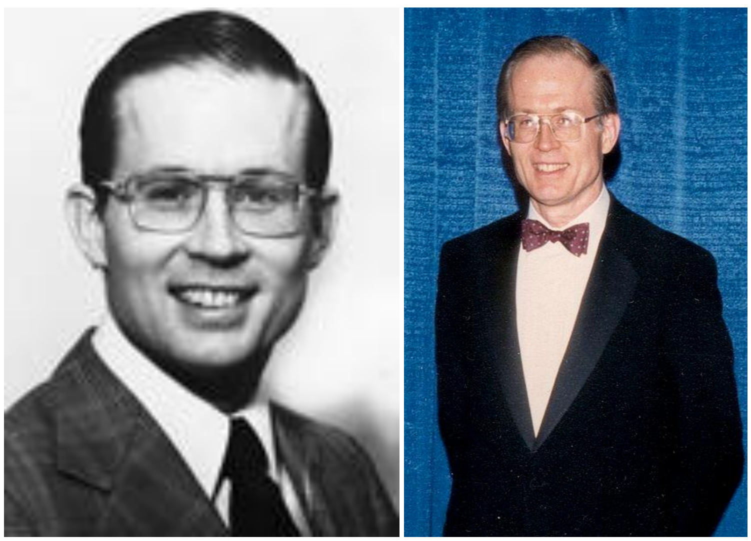
Minimizar $\sigma^2 = \sum_i \sum_j V_{ij} x_i x_j$ sujeito às restrições $\sum x_i = 1$, $\sum x_i \bar{r}_i = \bar{r}$ e $x_i \geq 0 \, \forall i$.

As restrições de desigualdade $x_i \geq 0$ complicam o problema matemático e tornam quase impossível encontrar solução analítica no caso geral.

Note que o desconhecimento dos retornos e da matriz de covariância no futuro não é exatamente uma falha do modelo de Markowitz. Se é possível prever, por qualquer critério, quem serão os retornos médios e a matriz de covariância, pode-se traçar a curva da fronteira eficiente. Entretanto, para ser útil o modelo tem que trazer alguma capacidade preditiva mesmo que seja em termos de probabilidades.

Modelo de Black: [http://en.wikipedia.org/wiki/Fischer_Black]

Fischer Black [1938-1995] graduou-se em física em 1959 por Harvard e defendeu seu doutorado, também em Harvard, em Matemática Aplicada em 1964. Não ganhou o prêmio Nobel pelo modelo Black&Scholes em 1997 porque morreu de câncer na garganta dois anos antes.



As hipóteses do modelo de Black são quase todas idênticas às do modelo de Markowitz com excessão da sexta, permitindo ao investidor ficar na posição short. A formulação matemática, então, do modelo de Black é dada por: encontrar o portfólio $\vec{X}$ que:

Minimiza $\sigma^2 = \sum_i \sum_j V_{ij} x_i x_j$ sujeito às restrições $\sum x_i = 1$, $\sum x_i \bar{r}_i = \bar{r}$.

Agora as restrições de desigualdade $x_i \geq 0$ desapareceram e o problema pode ser resolvido pela técnica dos multiplicadores de Lagrange. O modelo de Black admite solução analítica.

Nesse caso, o Lagrangeano do sistema é dado por:

$$\mathcal{L} = \sum_i \sum_j x_i x_j V_{ij} - \lambda_1 \left[\sum_i x_i \bar{r}_i - \bar{r} \right] - \lambda_2 \left[\sum_i x_i - 1 \right]$$

As condições de primeira ordem, $\frac{\partial \mathcal{L}}{\partial x_k} = 0$:

$$\frac{\partial}{\partial x_k} \mathcal{L} = \sum_i \sum_j x_i \frac{\partial x_j}{\partial x_k} V_{ij} + \sum_i \sum_j x_j \frac{\partial x_i}{\partial x_k} V_{ij} - \lambda_1 \left[\sum_i \frac{\partial x_i}{\partial x_k} \bar{r}_i \right] - \lambda_2 \left[\sum_i \frac{\partial x_i}{\partial x_k} \right]$$

$$\frac{\partial}{\partial x_k} \mathcal{L} = \sum_i \sum_j x_i \delta_{jk} V_{ij} + \sum_i \sum_j x_j \delta_{ik} V_{ij} - \lambda_1 \left[\sum_i \delta_{ik} \bar{r}_i \right] - \lambda_2 \left[\sum_i \delta_{ik} \right]$$

$$\frac{\partial}{\partial x_k} \mathcal{L} = \sum_i V_{ik} x_i + \sum_j V_{kj} x_j - \lambda_1 \bar{r}_k - \lambda_2 = \sum_i V_{ki} x_i + \sum_j V_{kj} x_j - \lambda_1 \bar{r}_k - \lambda_2$$

$$\frac{\partial}{\partial x_k} \mathcal{L} = 2 \sum_i V_{ki} x_i - \lambda_1 \bar{r}_k - \lambda_2 = 0 \quad \forall k$$

Essa condição pode ser escrita vetorialmente na forma $2V\vec{x} = \lambda_1 \vec{r} + \lambda_2$ ou ainda na forma matricial:

$$\begin{pmatrix} & & \\ & V_{ik} & \\ & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_A \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Já as duas restrições podem ser escritas em uma única forma matricial como:

$$\begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_A \end{pmatrix} = \begin{pmatrix} \bar{r} \\ 1 \end{pmatrix}$$

Sem saber os valores de λ_1 e λ_2 não saberemos o valor do portfólio. Nosso problema então é resolver as duas equações matriciais simultaneamente:

1. $V\,\vec{x} = \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$

2. $\begin{pmatrix} \overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_A \end{pmatrix} = \begin{pmatrix} \overline{r} \\ 1 \end{pmatrix}$

Substituindo o $\vec{x} = V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$ obtido na primeira equação na segunda equação obtemos $\begin{pmatrix} \overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \overline{r} \\ 1 \end{pmatrix}$, ou ainda $A \frac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \overline{r} \\ 1 \end{pmatrix}$, onde $A = \begin{pmatrix} \overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix}$ é uma matriz 2 x 2.

Propriedades da Matriz A

1. A é simétrica:

$$\overline{A} = \begin{pmatrix} \overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} \overline{V^{-1}} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} = \begin{pmatrix} \overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} = A.$$

Corolário: a matriz A^{-1} também é simétrica.

2. A é definida positiva:

Sabemos que V é simétrica e definida positiva, logo V^{-1} também é simétrica e definida positiva. Para isso precisamos encontrar o sinal de $q = \begin{pmatrix} h_1 & h_2 \end{pmatrix} A \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$. Usando a definição de A temos que:

$$q = \begin{pmatrix} h_1 & h_2 \end{pmatrix} \begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \dots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}.$$

Agora se definimos o vetor $\vec{k} = \begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ vemos que $\vec{k}' = \begin{pmatrix} h_1 & h_2 \end{pmatrix} \begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \dots & 1 \end{pmatrix}$, logo $q = \vec{k}' V^{-1} \vec{k}$. Mas a V^{-1} é definida positiva, logo $\begin{pmatrix} h_1 & h_2 \end{pmatrix} A \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} > 0$ para $\forall \vec{h} \neq \vec{0}$, ou seja, A é definida positiva.

Da equação $A\frac{1}{2}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=\begin{pmatrix}\bar{r}\\ 1\end{pmatrix}$ extraímos que $\frac{1}{2}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=A^{-1}\begin{pmatrix}\bar{r}\\ 1\end{pmatrix}$ que pode ser substituído na equação para o portfólio $\vec{X}=V^{-1}\begin{pmatrix}\bar{r}_1 & 1\\ \bar{r}_2 & 1\\ \vdots & \vdots\\ \bar{r}_A & 1\end{pmatrix}\frac{1}{2}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}$ e chegamos ao resultado do portfólio eficiente:

$$\vec{X}_{Black}=V^{-1}\begin{pmatrix}\bar{r}_1 & 1\\ \bar{r}_2 & 1\\ \vdots & \vdots\\ \bar{r}_A & 1\end{pmatrix}A^{-1}\begin{pmatrix}\bar{r}\\ 1\end{pmatrix}$$

Note que podemos encontrar o portfólio de Black para cada retorno médio desejado simplesmente como uma multiplicação de matrizes sobre o vetor $\begin{pmatrix}\bar{r}\\ 1\end{pmatrix}$, ou seja, $\vec{X}_{Black}=M\begin{pmatrix}\bar{r}\\ 1\end{pmatrix}$, onde $M=V^{-1}\begin{pmatrix}\bar{r}_1 & 1\\ \bar{r}_2 & 1\\ \vdots & \vdots\\ \bar{r}_A & 1\end{pmatrix}A^{-1}$. Só para checar as dimensões sabemos que um portfólio deve ser uma matriz $\vec{X}_{n\times 1}$ e que a matriz:

$$M_{n\times 2}=V_{n\times n}^{-1}\begin{pmatrix}\bar{r}_1 & 1\\ \bar{r}_2 & 1\\ \vdots & \vdots\\ \bar{r}_A & 1\end{pmatrix}_{n\times 2}A_{2\times 2}^{-1}$$

Logo $\vec{X}_{n\times 1}=M_{n\times 2}\begin{pmatrix}\bar{r}\\ 1\end{pmatrix}_{2\times 1}$ como deveria ser. Vamos escrever $M=\begin{pmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\\ \vdots & \vdots\\ \alpha_n & \beta_n\end{pmatrix}$ então podemos expressar o portfólio de Black como $\vec{X}_{Black}=\bar{r}\vec{\alpha}+\vec{\beta}$, ou seja:

$$\vec{X}_{Black} = \overline{r}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix} + \begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_n\end{pmatrix}$$

Com esse resultado podemos mostrar que uma combinação linear de dois portfólios eficientes é um portfólio eficiente. Sejam $\vec{X}_1 = \overline{r}_1\vec{\alpha} + \vec{\beta}$ e $\vec{X}_2 = \overline{r}_2\vec{\alpha} + \vec{\beta}$ dois portfólios eficientes de Black e vamos considerar o portfólio obtido com a combinação $\vec{X}_\lambda = \lambda\vec{X}_1 + (1-\lambda)\vec{X}_2$ desses dois portfólio. Usando a identidade acima temos que:

$$\vec{X}_\lambda = \lambda\left(\overline{r}_1\vec{\alpha} + \vec{\beta}\right) + (1-\lambda)\left(\overline{r}_2\vec{\alpha} + \vec{\beta}\right)$$

$$\vec{X}_\lambda = \left[\lambda\overline{r}_1 + (1-\lambda)\overline{r}_2\right]\vec{\alpha} + \left[\lambda + (1-\lambda)\right]\vec{\beta}$$

Ou seja $\vec{X}_\lambda = \overline{r}_\lambda\vec{\alpha} + \vec{\beta}$ com $\overline{r}_\lambda = \lambda\overline{r}_1 + (1-\lambda)\overline{r}_2$, também é um portfólio eficiente de Black.

Variância dos portfólios de Black:

Para calcular a variância dos portfólios de Black precisamos de $\sigma^2 = \overline{\vec{x}_{ef}} V\,\vec{x}_{ef}$. Mas $\vec{X}'_{Black} = \begin{pmatrix}\overline{r} & 1\end{pmatrix}\overline{A^{-1}}\begin{pmatrix}\overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A\\ 1 & 1 & \dots & 1\end{pmatrix}\overline{V^{-1}}$, que pode ser expresso, usando a simetria das matrizes A e V, como:

$$\vec{X}'_{Black} = \begin{pmatrix}\overline{r} & 1\end{pmatrix}A^{-1}\begin{pmatrix}\overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A\\ 1 & 1 & \dots & 1\end{pmatrix}V^{-1}$$

Dessa forma:

$$\sigma^2 = \begin{pmatrix}\overline{r} & 1\end{pmatrix}A^{-1}\begin{pmatrix}\overline{r}_1 & \overline{r}_2 & \cdots & \overline{r}_A\\ 1 & 1 & \dots & 1\end{pmatrix}V^{-1}VV^{-1}\begin{pmatrix}\overline{r}_1 & 1\\ \overline{r}_2 & 1\\ \vdots & \vdots\\ \overline{r}_A & 1\end{pmatrix}A^{-1}\begin{pmatrix}\overline{r}\\ 1\end{pmatrix}$$

Ou ainda:

$$\sigma^2 = \begin{pmatrix} \bar{r} & 1 \end{pmatrix} A^{-1} \begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \dots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix} A^{-1} \begin{pmatrix} \bar{r} \\ 1 \end{pmatrix}$$

Usando a definição de $A = \begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \dots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix}$ obtemos:

$$\sigma^2 = \begin{pmatrix} \bar{r} & 1 \end{pmatrix} A^{-1} A A^{-1} \begin{pmatrix} \bar{r} \\ 1 \end{pmatrix}.$$

Simplificando as multiplicações de matrizes por suas inversas chegamos ao resultado final bastante simples:

$$\sigma^2 = \begin{pmatrix} \bar{r} & 1 \end{pmatrix} A^{-1} \begin{pmatrix} \bar{r} \\ 1 \end{pmatrix}$$

Com esse resultado podemos mostrar que a fronteira eficiente de Black é uma hipérbole. Sabemos que $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ é uma matriz simétrica 2 x 2 definida positiva. Substituindo a inversa

$A^{-1} = \dfrac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$ na equação da variância obtemos:

$$\sigma^2 = \frac{1}{ac - b^2} \begin{pmatrix} \bar{r} & 1 \end{pmatrix} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \begin{pmatrix} \bar{r} \\ 1 \end{pmatrix} = \frac{1}{ac - b^2} \left(c\bar{r}^2 - 2b\bar{r} + a \right)$$

Completando quadrado nos termos com $\bar{r}$ somando e subtraindo $\dfrac{b^2}{c}$ re-escrevemos a equação como:

$$\sigma^2 = \frac{1}{ac-b^2}\left[c\left(\bar{r}-\frac{b}{c}\right)^2 + \frac{ac-b^2}{c}\right]$$

Com um pouco de álgebra simples podemos expressar essa curva como:

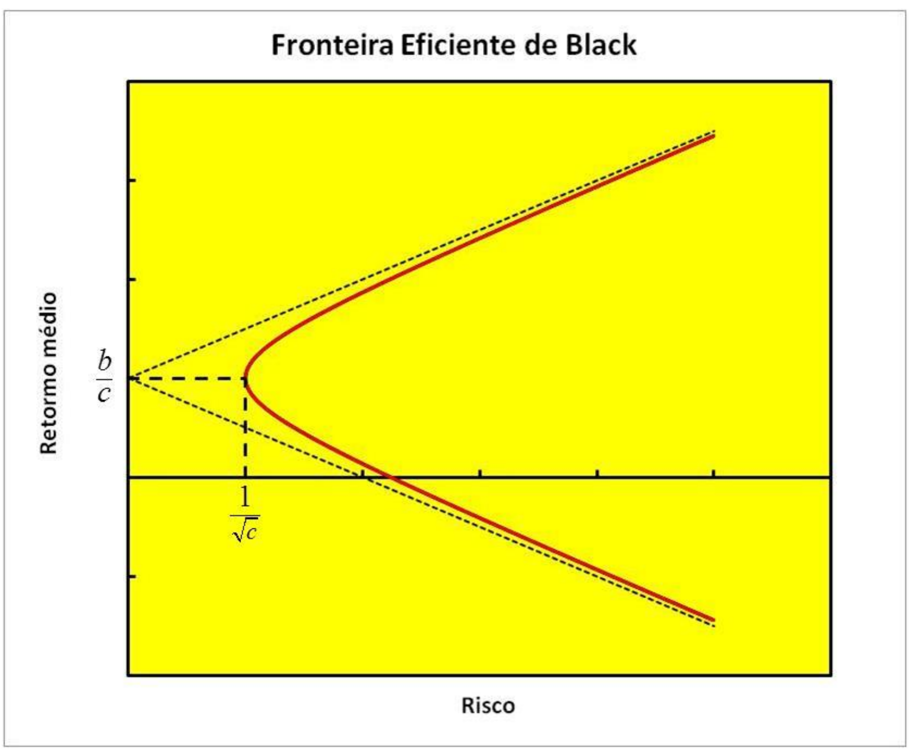
$$\frac{\sigma^2}{\left(\frac{1}{\sqrt{c}}\right)^2} - \frac{\left(\bar{r}-\frac{b}{c}\right)^2}{\left(\frac{\sqrt{ac-b^2}}{c}\right)^2} = 1$$

Que é equação do lugar geométrico de uma hipérbole do tipo:

$$\frac{x^2}{\alpha^2} - \frac{\left(y-y_o\right)^2}{\beta^2} = 1$$

Note então que $y = y_o \pm \frac{\beta}{\alpha}\sqrt{x^2-\alpha^2}$ formam os dois braços do ramo positivo da hipérbole.

Notamos que só existe $y \in \mathbb{R}$ se $x \geq \alpha$ e que se $x >> \alpha$ caímos nas assíntotas inclinadas $y = y_o \pm \frac{\beta}{\alpha}x$. A figura xx abaixo mostra a hipérbole obtida para a fronteira eficiente de Black.



Para traçar a curva de uma hipérbole é melhor usar as equações paramétricas: $x = \alpha \cosh t$ e $y = y_o + \beta \sinh t$. Note que $\frac{x^2}{\alpha^2} = \cosh^2 t$ e $\frac{(y - y_o)^2}{\beta^2} = \sinh^2 t$ logo $\frac{x^2}{\alpha^2} - \frac{(y - y_o)^2}{\beta^2} = \cosh^2 t - \sinh^2 t$. Mas $\cosh^2 t - \sinh^2 t = 1$, pois:

$$\cosh^2 t - \sinh^2 t = \frac{\left(e^t + e^{-t}\right)^2}{4} - \frac{\left(e^t - e^{-t}\right)^2}{4} = \frac{e^{2t} + 2 + e^{-2t}}{4} - \frac{e^{2t} - 2 + e^{-2t}}{4} = \frac{4}{4} = 1$$

e reobtemos a equação do lugar geométrico: $\frac{x^2}{\alpha^2} - \frac{(y - y_o)^2}{\beta^2} = 1$ CQD.

Portfólio do vértice: no vértice $\overline{r} = \frac{b}{c}$ então:

$$\vec{x}_{vert} = V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \begin{pmatrix} \frac{b}{c} \\ 1 \end{pmatrix} = V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \frac{1}{ac - b^2} \begin{pmatrix} c\frac{b}{c} - b \\ -\frac{b^2}{c} + a \end{pmatrix}$$

$$\vec{x}_{vert} = V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \frac{1}{ac - b^2} \begin{pmatrix} 0 \\ \frac{ac - b^2}{c} \end{pmatrix}$$

$$\vec{x}_{vert} = \frac{1}{c}V^{-1}\begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{c}V^{-1}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\vec{x}_{vert} = \frac{1}{c}V^{-1}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

O portfólio do vértice poderia ser colocado como o seguinte problema:

Portfólio de Mínima Variância

O modelo de mínima variância é expresso, então, como o problema de otimização dado por: Minimizar $\sigma^2 = \sum_i \sum_j V_{ij} x_i x_j$ sujeito à restrição $\sum x_i = 1$. O problema é resolvido do mesmo modo que o de Black.

O Lagrangeano do sistema é dado por: $\mathcal{L} = \sum_i \sum_j x_i x_j V_{ij} - \lambda \left[\sum_i x_i - 1 \right]$

As condições de primeira ordem $\frac{\partial \mathcal{L}}{\partial x_k} = 0$ podem ser escritas na forma matricial como:

$$\begin{pmatrix} & & \\ & V_{ik} & \\ & & \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_A \end{pmatrix} = \frac{1}{2}\lambda\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ ou } \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_A \end{pmatrix} = \frac{1}{2}V^{-1}\lambda\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Para determinar o valor do portfólio calculamos λ . A restrição pode ser escrita como:

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_A \end{pmatrix} = 1$$

Logo:

$$\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_A\end{pmatrix}=\frac{\lambda}{2}\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}V^{-1}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}=1$$

Então $\dfrac{\lambda}{2}=\dfrac{1}{\left[\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}V^{-1}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\right]}$.

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_A\end{pmatrix}=\frac{1}{\left[\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}V^{-1}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\right]}V^{-1}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$$

Agora notamos que na matriz $A=\begin{pmatrix}a & b\\ b & c\end{pmatrix}=\begin{pmatrix}\bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A\\ 1 & 1 & \dots & 1\end{pmatrix}V^{-1}\begin{pmatrix}\bar{r}_1 & 1\\ \bar{r}_2 & 1\\ \vdots & \vdots\\ \bar{r}_A & 1\end{pmatrix}$ o elemento $c=A_{22}$

é dado exatamente por: $A_{22}=\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix}V^{-1}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$, e caímos de volta no resultado:

$$\vec{x}_{vert}=\frac{1}{c}V^{-1}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$$

Podemos calcular cada elemento através de:

$$c = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_i \sum_j V_{ij}^{-1}$$

$V^{-1} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_j V_{ij}^{-1}$ logo:

$$x_i = \frac{\sum_j V_{ij}^{-1}}{\sum_i \sum_j V_{ij}^{-1}}$$

Modelo de Tobin. [http://en.wikipedia.org/wiki/James_Tobin]

James Tobin [1918 – 2002] formou-se em economia por Harvard em uma atmosfera pioneira em economia do qual sairam vários prêmios Nobel, incluindo o de Tobin em 1981. Iniciou seu trabalho em pesquisas Keynesianas mas trouxe contribuições importantes nas áreas de investimento e mercados financeiros, políticas fiscal e monetárias, econometria onde desenvolveu o modelo TOBIT. Fora da economia sua proposta de um imposto sobre transações internacionais, denominado Tobin Tax, para reduzir a especulação nos mercados internacionais de câmbio que considerava perigosa.

Tobin introduziu um ativo de renda fixa, que nas hipóteses de mercado ideal é único, ou seja, todos os ativos de renda fixa apresentam o mesmo retorno. No modelo de Tobin o investidor pode ficar nas posições long e short no ativo de renda fixa mas apenas na posição long nos ativos arriscados. A variância da renda fixa é nula, i.e., $\sigma_f^2 = 0$ e, obviamente, $\overline{r_f} = r_f$.

Considere um portfólio com ativos arriscados apenas com rendimento r_p e variância σ_p^2. Investindo uma fração x na renda fixa e $(1-x) > 0$ no portfólio arriscado, o novo portofólio apresenta retornos dados por:

$$r = xr_f + (1-x)r_p$$
$$\overline{r} = xr_f + (1-x)\overline{r}_p$$

Então $r - \overline{r} = (1-x)(r_p - \overline{r}_p)$ e $(r - \overline{r})^2 = (1-x)^2 (r_p - \overline{r}_p)^2$ que leva a:

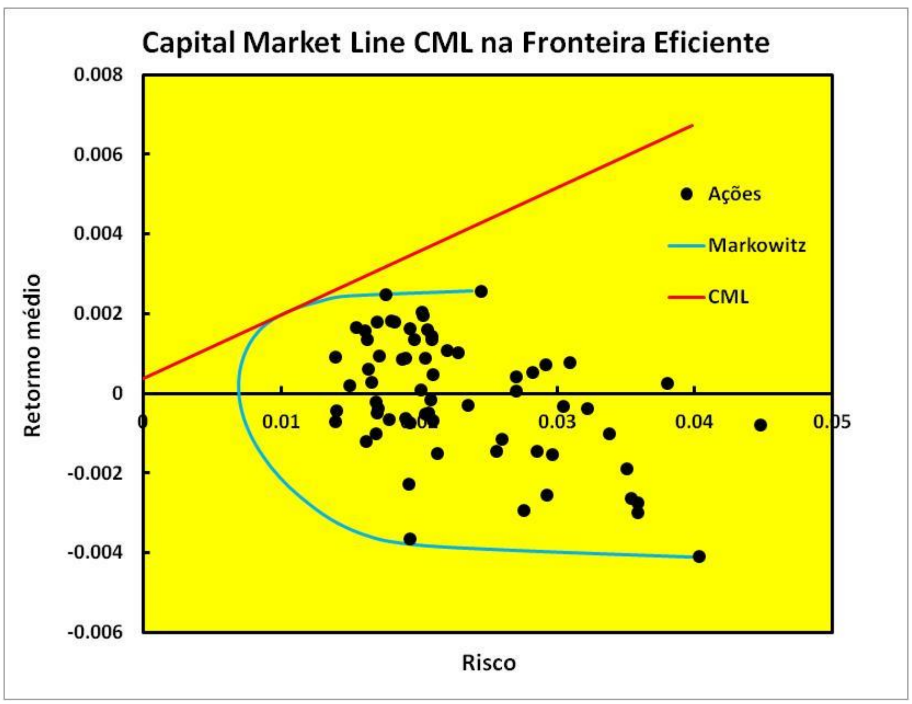
$$\sigma_r^2 = E\left[(r - \overline{r})^2\right] = (1-x)^2 E\left[(r_p - \overline{r}_p)^2\right]$$

$\sigma_r^2 = (1-x)^2 \sigma_p^2$ ou, sabendo que $(1-x) > 0$:

$\sigma_r = (1-x)\sigma_p$.

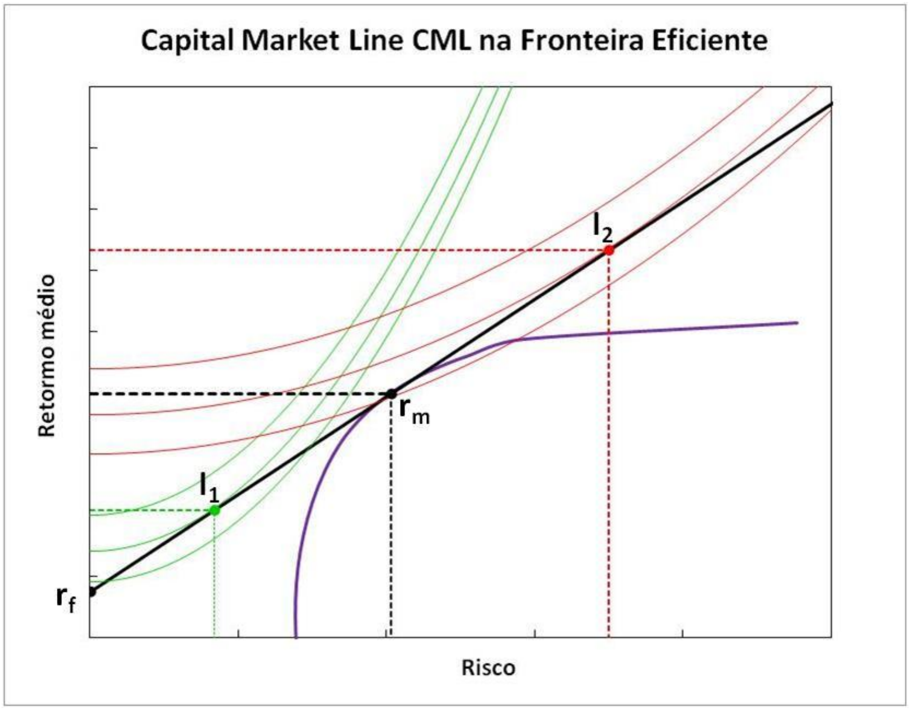
Dessa equação extraímos que $\left(1-x\right)=\frac{\sigma_r}{\sigma_p}$, $x=1-\frac{\sigma_r}{\sigma_p}$.

Dessa forma $\overline{r}=\left(1-\frac{\sigma_r}{\sigma_p}\right)r_f+\frac{\sigma_r}{\sigma_p}\overline{r}_p$ ou seja, $\overline{r}=r_f+\frac{\left(\overline{r}_p-r_f\right)}{\sigma_p}\sigma_r$. Na curva retorno-risco isso representa uma reta que passa pelo ponto $\left(0,r_f\right)$ com inclinação $\frac{\left(\overline{r}_p-r_f\right)}{\sigma_p}$. Com $0\leq x\leq 1$ obtém-se um segmento de reta ligando os pontos $\left(0,r_f\right)$ e $\left(\sigma_p,r_p\right)$. Com $x<0$ podemos continuar a reta indefinidamente, como mostra a figura xxx abaixo.



A nova fronteira eficiente após a introdução da renda fixa, então, é a reta tangente que passa pela renda fixa e tangencia a fronteira de Markowitz. Qualquer ponto dessa fronteira pode ser alcançado com apenas dois portfólios: a renda fixa e o portfólio no ponto de tangência, que chamaremos de portfólio de mercado $\vec{X}_m$, com retorno $\overline{r}_m$ e desvio padrão σ_m. Essa reta é chamada de **CAPITAL MARKET LINE**.

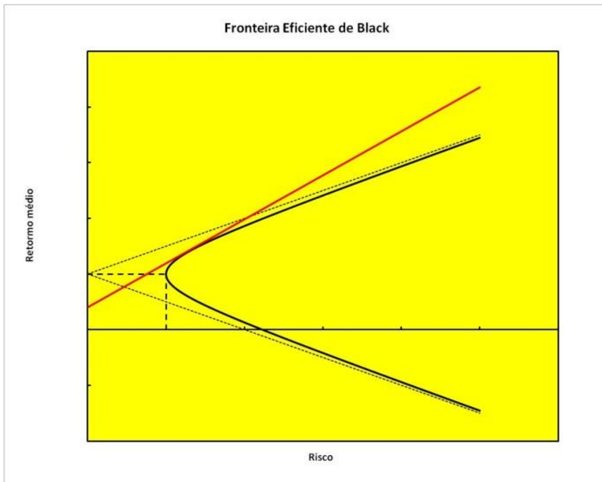
Se todos os investidores usam as mesmas avaliações de retorno-risco baseados no desempenho histórico a CML será a mesma para todo mundo. Nesse caso supondo que todos os investidores são avessos ao risco a única decisão do investidor é escolher o ponto da CML para seu investimento, ou seja, de que forma financiará o portfólio de mercado. Eles escolherão seus portfólios ao longo da CML de acordo com suas preferências, como mostra a figura xxx abaixo com dois investidores, I_1 e I_2.

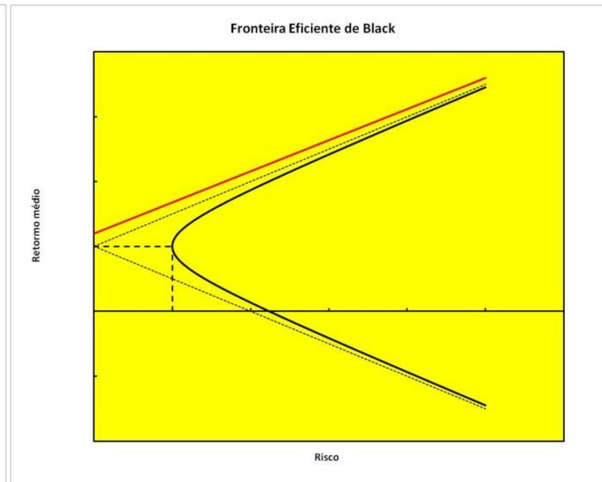


Nesse caso o investidor 1 é chamado de CONSERVADOR, pois comprou quantidades positivos de renda fixa e do portfólio de mercado, ou seja, financiou o portfólio de mercado com capital próprio. Se o ativo arriscado sair perdendo ele ganha pelo menos a renda fixa. Já o investidor 2 é chamado de AGRESSIVO, pois pegou dinheiro emprestado na renda fixa para comprar mais do portfólio de mercado. Financiou o portfólio de mercado com empréstimo, está alavancado. Se houver prejuízo no ativo arriscado ele perde mais do que o retorno negativo, pois ainda terá que cobrir o empréstimo na taxa da renda fixa.

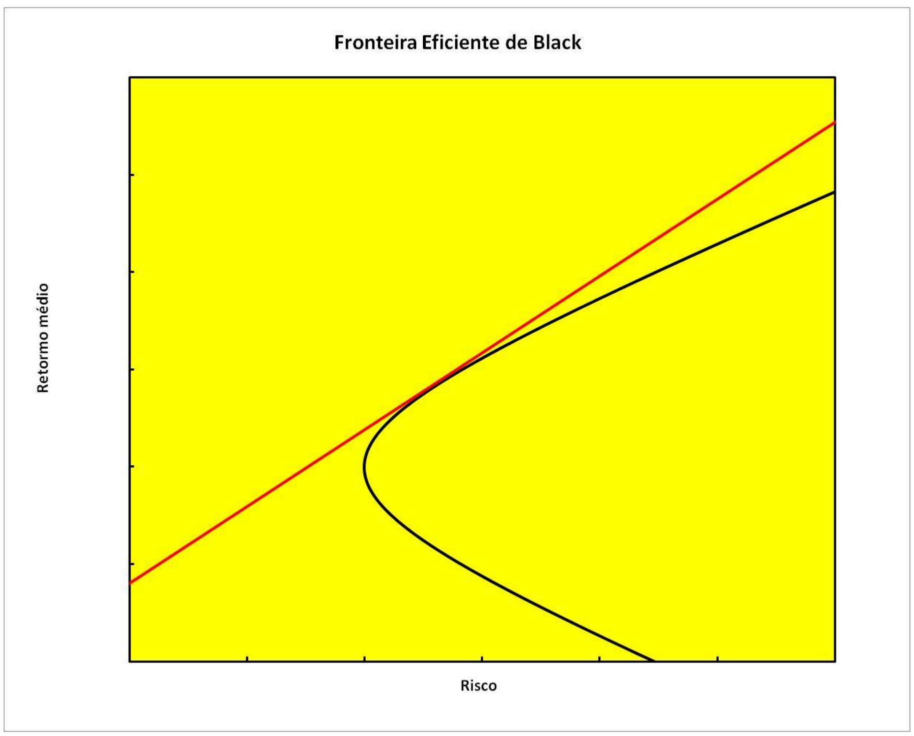
Modelo de Tobin/Black.

Nesse modelo utiliza-se a fronteira eficiente de Black em lugar de Markowitz. Note que só haverá solução para o portfólio de mercado no modelo de Black se o retorno da renda fixa estiver abaixo do vértice da fronteira eficiente, i.e., $r_f < \frac{b}{c}$, caso contrário a reta tangente se torna paralela à assintóta inclinada da hipérbole.





No caso de Tobin/Black é possível encontrar uma expressão analítica para a CML.



A expressão da reta é dada por: $\overline{r} = r_f + \dfrac{d\,\overline{r}}{d\sigma}\sigma$ enquanto a expressão da hipérbole é dada por: $c\sigma^2 - \dfrac{c^2}{ac-b^2}\left(\overline{r} - \dfrac{b}{c}\right)^2 = 1$. Derivando a hipérbole implicitamente temos:

$$2c\sigma - \frac{c^2}{ac-b^2} 2\left(\bar{r} - \frac{b}{c}\right)\frac{d\,\bar{r}}{d\sigma} = 0$$

Logo $\frac{d\,\bar{r}}{d\sigma} = \frac{ac-b^2}{c} \frac{\sigma}{\left(\bar{r} - b/c\right)}$

Substituindo na equação da reta:

$$\bar{r} = r_f + \frac{ac-b^2}{c^2} \frac{c\sigma^2}{\left(\bar{r} - b/c\right)}$$

Substituindo $c\sigma^2$ da hipérbole temos:

$$\bar{r} = r_f + \frac{ac-b^2}{c^2} \frac{1}{\left(\bar{r} - b/c\right)}\left[1 + \frac{c^2}{ac-b^2}\left(\bar{r} - b/c\right)^2\right]$$

Logo:

$$\bar{r} = r_f + \frac{ac-b^2}{c^2} \frac{1}{\left(\bar{r} - b/c\right)} + \bar{r} - b/c$$

Ou seja:

$$\bar{r}_m = \frac{b}{c} + \frac{ac-b^2}{c^2} \frac{1}{\left(b/c - r_f\right)}$$

Já o σ_m pode ser calculado por $c\sigma_m^2 = 1 + \frac{c^2}{ac-b^2}\left(\bar{r}_m - \frac{b}{c}\right)^2$, ou seja, será dado por:

$$c\sigma_m^2 = 1 + \frac{c^2}{ac-b^2}\left(\frac{ac-b^2}{c^2}\right)^2 \frac{1}{\left(b/c - r_f\right)^2}$$

Ou ainda:

$$\sigma_m = \frac{1}{\sqrt{c}}\sqrt{1 + \frac{ac-b^2}{c^2}\frac{1}{\left(b/c - r_f\right)^2}}$$

A inclinação da reta é dada por:

$$\frac{\bar{r}_m - r_f}{\sigma_m} = \frac{\left(b/c - r_f\right) + \frac{ac-b^2}{c^2}\frac{1}{\left(b/c - r_f\right)}}{\frac{1}{\sqrt{c}}\sqrt{1 + \frac{ac-b^2}{c^2}\frac{1}{\left(b/c - r_f\right)^2}}}$$

Ou seja:

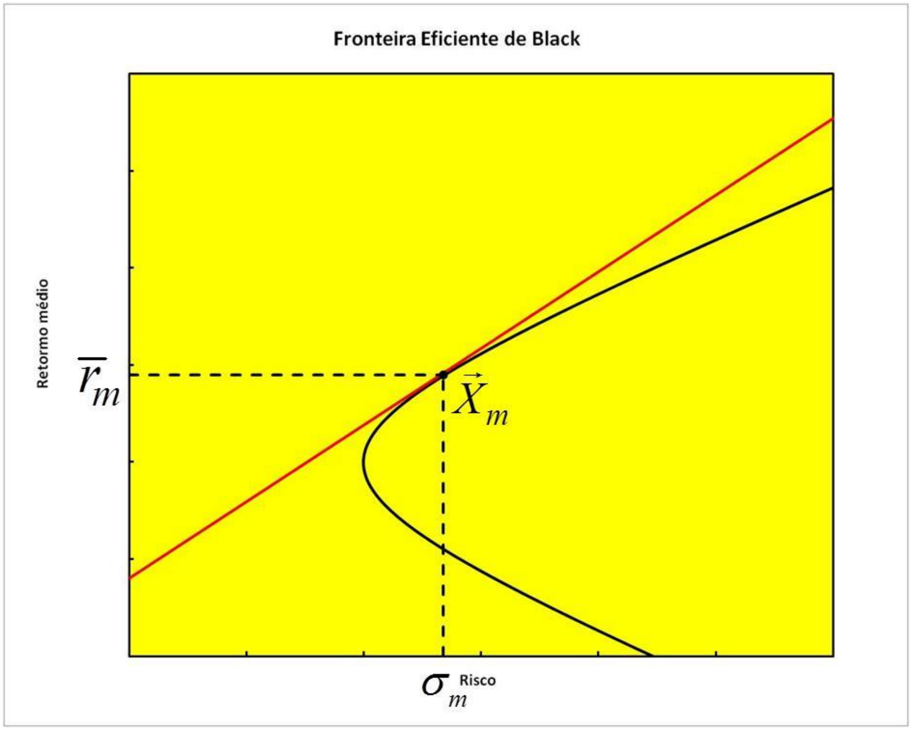
$$\frac{d\,\bar{r}}{d\sigma} = \sqrt{c}\left(b/c - r_f\right)\sqrt{1 + \frac{ac-b^2}{c^2}\frac{1}{\left(b/c - r_f\right)^2}}$$

Finalmente:

$$\frac{d\,\bar{r}}{d\sigma} = \sqrt{c}\sqrt{\left(b/c - r_f\right)^2 + \frac{ac-b^2}{c^2}}$$

O portofólio de mercado é calculado com as vetores $\vec{\alpha}$ e $\vec{\beta}$ dado por:

$$\vec{X}_m = \vec{\alpha}\,\bar{r}_m + \vec{\beta}$$



Modelo CAPM [Capital asset pricing model]:

Este modelo foi desenvolvido por Jack Treynor (1961, 1962), William Sharpe (1964), John Lintner (1965) e Jan Mossin (1966) independentemente, sobre as bases da MPT de Markowitz.

Suposições do modelo:

1. Todos os investidores são avessos ao risco d maximizam suas utilidades esperadas em um período.
2. Escolhem seus portfólios apenas em termos de retorno médio e risco.
3. Existem $\sigma \neq 0$ e $\overline{r}$ para qualquer portfólio.
4. Todos os títulos são infinitamente divisíveis.
5. Investidores são price-takers.

As suposições até esse ponto são suficientes para garantir que todos os investidores escolham seus portfólios na fronteira eficiente. Além disso, existem mais suposições:

6. Existem taxas de empréstimo e crédito sem risco.

7. Mercado de capitais é pefeito, i.e.:

7.1. Toda informação é grátis e instantaneamente disponível à todos

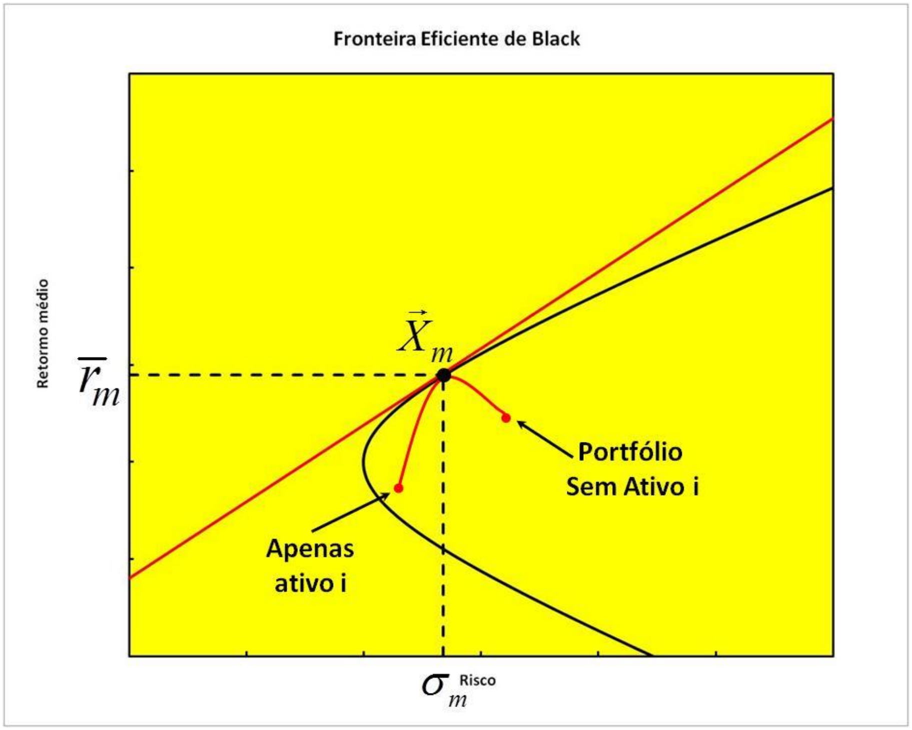
7.2. Não há limites, ou margens, para as quantidades que os investidores podem comprar ou vender.

7.3. Taxa de empréstimo = taxa de crédito.

8. Os investidores possuem expectativas homogêneas, todos têm o mesmo horizonte de tempo, ou seja, o período de todos é o mesmo, e utilizam o mesmo conjunto dos vetores r médio e matriz de covariância.
9. Todos os ativos são negociáveis, incluindo capital humano.

SML: Security Market Line.

Vamos tomar um portfólio composto pela combinação convexa do portfólio de mercado e o ativo i isolado, ou seja, $\vec{X}_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ e $\vec{X}_m$ de modo que $\vec{X}_\lambda = \lambda \vec{X}_i + (1-\lambda)\vec{X}_m$. Para $\lambda > 0$ estamos aumentando a participação do iésimo ativo no portfólio de mercado. Se $\lambda = 1$ o portfólio só tem o iésimo ativo, e se $\lambda = 0$ adquirimos exatamente o portfólio de mercado. Também podemos fazer $\lambda < 0$ até zerar a participação do iésimo ativo no portfólio final. Claramente a curva obtida variando λ está incluída na fornteira eficiente e é tangente à mesma quando $\lambda = 0$ como mostra a figura xxx.



Agora para esse portfólio:

$$r = \lambda r_i + (1-\lambda) r_m$$
$$\overline{r} = \lambda \overline{r}_i + (1-\lambda) \overline{r}_m$$

Logo $(r-\overline{r}) = \lambda(r_i - \overline{r}_i) + (1-\lambda)(r_m - \overline{r}_m)$ e:

$$(r-\overline{r})^2 = \lambda^2 (r_i - \overline{r}_i)^2 + (1-\lambda)^2 (r_m - \overline{r}_m)^2 + 2(1-\lambda)\lambda(r_i - \overline{r}_i)(r_m - \overline{r}_m)$$

Então:

$$E\left[(r-\overline{r})^2\right] = \lambda^2 E\left[(r_i - \overline{r}_i)^2\right] + (1-\lambda)^2 E\left[(r_m - \overline{r}_m)^2\right] + 2(1-\lambda)\lambda E\left[(r_i - \overline{r}_i)(r_m - \overline{r}_m)\right]$$

Que significa que:

$$\sigma^2 = \lambda^2 \sigma_i^2 + (1-\lambda)^2 \sigma_m^2 + 2\lambda(1-\lambda)V_{im}$$

onde $V_{im} = E\left[(r_i - \overline{r}_i)(r_m - \overline{r}_m)\right] = \operatorname{cov}(r_i, r_m)$

Assim temos duas equações:

$$\bar{r} = \lambda \bar{r}_i + (1-\lambda)\bar{r}_m$$

$$\sigma^2 = \lambda^2 \sigma_i^2 + (1-\lambda)^2 \sigma_m^2 + 2\lambda(1-\lambda)V_{im}$$

Das quais desejamos extrair $\frac{d\,\bar{r}}{d\sigma}$ e obrigar que a inclinação seja a mesma da CML para $\lambda = 0$. Mlehor expressar a derivada em termos do parâmetro λ na forma $\frac{d\,\bar{r}}{d\sigma} = \frac{\left(\frac{d\,\bar{r}}{d\lambda}\right)}{\left(\frac{d\sigma}{d\lambda}\right)}$. A derivada do numerador é imediata: $\frac{d\,\bar{r}}{d\lambda} = \bar{r}_i - \bar{r}_m$. A do denominador pode ser obtida por derivação implícita:

$$2\sigma \frac{d\sigma}{d\lambda} = 2\lambda\sigma_i^2 - 2(1-\lambda)\sigma_m^2 + 2(1-\lambda)V_{im} - 2\lambda V_{im}$$

Ou seja:

$$\sigma \frac{d\sigma}{d\lambda} = \lambda\sigma_i^2 - (1-\lambda)\sigma_m^2 + (1-\lambda)V_{im} - \lambda V_{im}$$

Para $\lambda = 0$ então $\sigma = \sigma_m$:

$$\sigma_m \frac{d\sigma_m}{d\lambda} = V_{im} - \sigma_m^2$$

Logo:

$$\frac{d\sigma_m}{d\lambda} = \frac{V_{im} - \sigma_m^2}{\sigma_m}$$

Assim $\frac{d\,\bar{r}}{d\sigma} = \frac{(\bar{r}_i - \bar{r}_m)}{(V_{im} - \sigma_m^2)}\sigma_m$ deve ser igual à inclinação da CML $\frac{d\,\bar{r}}{d\sigma} = \frac{\bar{r}_m - r_f}{\sigma_m}$, o que significa que:

$\frac{(\bar{r}_i - \bar{r}_m)}{(V_{im} - \sigma_m^2)}\sigma_m = \frac{\bar{r}_m - r_f}{\sigma_m}$ ou seja, $\bar{r}_i - \bar{r}_m = \frac{\bar{r}_m - r_f}{\sigma_m^2}(V_{im} - \sigma_m^2)$, ou ainda:

$$\overline{r}_i - \overline{r}_m = r_f - \overline{r}_m + \frac{\overline{r}_m - r_f}{\sigma_m^2} V_{im}$$

Finalmente: $\overline{r}_i = r_f + \dfrac{\overline{r}_m - r_f}{\sigma_m^2} V_{im}$

A qual pode ser escrita na forma $\overline{r}_i = r_f + \beta V_{im}$. Essa é a equação da reta chamada Security Market Line e afirma que o excesso de retorno em relação à renda fixa $\overline{r}_i - r_f$ é diretamente proporcional à covariância do ativo i com o portfólio de mercado. Note que o modelo CAPM não exige que se use o modelo de Black, ou mesmo de Markowitz. Pode ser usado com qualquer modelo que usa o desvio padrão como uma proxy do risco.

Entretanto, o CAPM inclui suposições muito fortes. Considere o késimo investidor e seu gasto no iésimo ativo: $x_{ik} = \dfrac{W_{ki}}{W_k}$, logo $W_{ki} = x_{ik} W_k$. Obviamente $\sum_i W_{ki} = W_k \sum_i x_{ik} = W_k$ pois $\sum_i x_{ik} = 1$. A suposição do CAPM é que todos os agentes têm acesso à mesma fronteira eficiente, logo todos eles utilizarão o mesmo portfólio de mercado o que significa que $x_{ik} = x_i$, independente de k, será o mesmo para todos os agentes. Mas W_k muda de agente para agente. Somando todos os gastos no iésimo ativo:

$$\sum_k W_{ki} = \sum_k x_{ik} W_k = x_i \sum_k W_k$$

Portanto:

$$x_i = \frac{\sum_k W_{ki}}{\sum_k W_k} = \frac{\text{soma de todos os investimentos em i}}{\text{soma total de todos os investimentos}}$$

Isso significa que o portfólio de mercado deve incluir quantidades positivas não nulas de todos os ativos negociados no mercado. A dificuldade com os modelos que usamos vem de: (1) no portfólio de Markowitz a participação de muitos ativos é nula e (2) no de Black várias são negativas. Os dados para o cálculo de x_i acima são observáveis – as bolsas trazem a informação sobre volume de negócios em cada uma das ações e de seus preços. Dessa forma seria possível calcular um portfólio de mercado e acompanhar seu desempenho no tempo, calcular a covariância com cada um dos ativos e observar a SML. Dados sobre portfólio de mercado são disponibilizados nas bolsas americanas, pelo menos. Proxys desse portfólio seriam os índices das bolsas, como o IBOVESPA, Dow Jones, etc. Entretanto, os construtores do modelo tornaram a verificação muito mais complicada, porque o modelo supõe todos os possíveis ativos como bens imóveis, commodities, capital humano etc, sobre os quais não é possível traçar um curva de fronteira eficiente.

Se calculamos o portfólio de mercado como o portfólio que tangencia a fronteira eficiente a SML é perfeita. Entretanto, usando uma proxy do portfólio de mercado os resultados não são nada bons.

Alternativas à Markowitz/Black/Tobin/CAPM:

Risk Budgeting:

Sabemos que $\sigma^2 = \sum_i \sum_j V_{ij} x_i x_j$ logo $2\sigma \frac{\partial \sigma}{\partial x_k} = \sum_i \sum_j V_{ij} x_i \frac{\partial x_j}{\partial x_k} + \sum_i \sum_j V_{ij} x_j \frac{\partial x_i}{\partial x_k}$, ou seja:

$$2\sigma \frac{\partial \sigma}{\partial x_k} = \sum_i \sum_j V_{ij} x_i \delta_{jk} + \sum_i \sum_j V_{ij} x_j \delta_{ik}$$

$$2\sigma \frac{\partial \sigma}{\partial x_k} = \sum_i V_{ik} x_i + \sum_j V_{kj} x_j = 2\sum_i V_{ki} x_i$$

Portanto:

$$\frac{\partial \sigma}{\partial x_k} = \frac{1}{\sigma} \sum_i V_{ki} x_i$$

Agora $\sum_k x_k \frac{\partial \sigma}{\partial x_k} = \frac{1}{\sigma} \sum_k \sum_i V_{ki} x_i = \frac{1}{\sigma} \sigma^2 = \sigma$ portanto:

$$\sum_k \left(\frac{x_k}{\sigma} \frac{\partial \sigma}{\partial x_k} \right) = 1$$

O que essa expressão está afirmando é que a soma das elasticidades x do risco vale 1. Elasticidade x de y é definida como $e_{x-y} = \frac{x}{y} \frac{\partial y}{\partial x} = \frac{\left(\frac{dy}{y} \right)}{\left(\frac{dx}{x} \right)}$ é uma grandeza adimensional que fornece a informação sobre quanto é a variação relativa, percentual, em y devido a uma variação relativa, percentual, em x. Note que $e_{x_k-\sigma} = \frac{x_k}{\sigma}\frac{\partial \sigma}{\partial x_k}$ e que $\sum_k e_{x_k-\sigma} = 1$.

Podemos calcular o vetor $\frac{1}{\sigma}\frac{\partial \sigma}{\partial x_k}$ diretamente através de operação matricial da forma: $\overrightarrow{\left(\frac{1}{\sigma}\frac{\partial \sigma}{\partial x_k}\right)} = \frac{1}{\sigma^2}V\vec{X}$. Daí cada elasticidade é dada por $x_k\left(\frac{1}{\sigma}\frac{\partial \sigma}{\partial x_k}\right)$. A idéia do risk-budgeting é montar o portfólio impondo que a elasticidade x do risco de cada ativo não seja nem muito alta nem baixa demais. Por exemplo, podemos impor que $\frac{1}{10N} \le e_{x_k-\sigma} \le \frac{1}{N}$, onde N é o número de ativos que estamos utilizando. Esse procedimento impede que a participação de qualquer ativo seja nula e que se concentre o risco em poucos ativos. A implementação em uma planilha excel envolve incluir mais N restrições na macro em VBA.

Modelo de SEMI-VARIÂNCIA:

No modelo de Markowitz/Black a proxy do risco foi a variância. Entretanto, risco para o investidor é risco de perder e não de ganhar. A variância inclui variações para baixo e para cima. Gostaríamos de minimizar apenas a parte negativa dos retornos, e não a parte positiva. Todos desejamos correr o risco de ganhar mais. Se a distribuição de variações for simétrica, a semivariância negativa será igual à positiva e tanto faz minimizar a variância quanto a semivariância. Entretanto, isso não é mais verdade para distribuições assimétricas.

A matemática do modelo de semivariância é:

Minimizar $\sigma_{semi} = \frac{1}{n}\sum_t \left(r_t - r_{\text{target}}\right)^2 H\left(r_{\text{target}} - r_t\right)$,

onde $H\left(r_{\text{target}} - r_t\right) = \begin{cases} 0 & se\; r > r_{\text{target}} \\ 1 & se\; r < r_{\text{target}} \end{cases}$, sujeito às restrições: $\sum_i x_i \overline{r_i} = r_{\text{target}}$; $\sum_i x_i = 1$ e $x_i \ge 0 \quad \forall i$.

A implementação no Excel é direta. Chutamos valores para $\left(x_1, x_2, \cdots, x_N\right)$ com os quais calculamos $r_t = \sum_i x_i r_i$. Escolhemos r_{target} e sabemos, para cada t, calcular $\left(r_t - r_{\text{target}}\right)^2 H\left(r_{\text{target}} - r_t\right)$. Somamos então esse valores em uma célula alvo e usamos o SOLVER com as restrições para minimizar esse valor encontrando o portfólio que minimiza a semivariância. Repetimos o processo para cada r_{target}. Note que podemos usar também modelo de semivariância de Black.

Single Index Model [SIM]:

A idéia desse modelo é que cada ativo tem uma dependência com algum índice do mercado como um todo e um comportamento independente. Ou seja, a covariância entre os ativos advém do fato de que todos eles dependem do mercado. Nesse modelo então:

$$r_i = \alpha_i + \beta_i r_m + e_i$$

Onde o iésimo erro e_i tem as seguintes propriedades: $E\left[e_i\right] = 0$, $E\left[e_i e_{j \neq i}\right] = 0$ e $E\left[e_i^2\right] = \sigma_{io}^2$.

Neste caos notamos que $\overline{r}_i = \alpha_i + \beta_i \overline{r}_m$ logo:

$$r_i - \overline{r}_i = \beta_i\left(r_m - \overline{r}_m\right) + e_i$$
$$r_j - \overline{r}_j = \beta_j\left(r_m - \overline{r}_m\right) + e_j$$

Então

$$\left(r_i - \overline{r}_i\right)\left(r_j - \overline{r}_j\right) = \beta_i \beta_j \left(r_m - \overline{r}_m\right)^2 + \beta_i\left(r_m - \overline{r}_m\right)e_j + \beta_j\left(r_m - \overline{r}_m\right)e_i + e_i e_j$$

Logo:

$$E\left[\left(r_i - \overline{r}_i\right)\left(r_j - \overline{r}_j\right)\right] = \beta_i \beta_j E\left[\left(r_m - \overline{r}_m\right)^2\right] + \beta_i E\left[\left(r_m - \overline{r}_m\right)e_j\right] + \beta_j E\left[\left(r_m - \overline{r}_m\right)e_i\right] + E\left[e_i e_j\right]$$

$\operatorname{cov}\left(r_i, r_j\right) = \beta_i \beta_j \sigma_m^2 + \sigma_{io}^2 \delta_{ij}$ de onde tiramos que:

$$V_{ij} = \begin{cases} \beta_i \beta_j \sigma_m^2 & j \neq i \\ \beta_i^2 \sigma_m^2 + \sigma_{io}^2 & j = i \end{cases}$$

Que pode ser re-escrito como:

$$V_{ij} = \begin{cases} \beta_i \beta_j \sigma_m^2 & j \neq i \\ \sigma_i^2 & j = i \end{cases}$$

A proxy para o mercado pode ser um índice como IBOVESPA, Dow Jones etc. Encontramos os $\beta's$ através de uma regressão de r_i com r_m. Na realidade o excel já permite calcular apenas o coeficiente angular da regressão com a função SLOPE. A diagonal da matriz de covariância encontramos usando a variância de cada ação, e calculamos a matriz de covariância usando a variância da proxy do mercado multiplicando pelos devidos $\beta's$. Dessa forma podemos re-obter a matriz de variância-covariância a ser usada em um modelo Markowitz/Black. Trata-se então de uma forma de tentar limpar a matriz de variância-covariância. Note que o portfólio de mercado poderia também ser o primeiro portfólio que encontramos na diagonalização da matriz de covariância.

Risco de Mercado e Risco Diversificável:

Usando o SIM para um portfólio qualquer temos que:

$$r = \sum_i x_i r_i$$

$$r = \sum_i x_i \left(\alpha_i + \beta_i r_m + e_i \right)$$

$$r = \sum_i x_i \alpha_i + \sum_i x_i e_i + r_m \sum_i x_i \beta_i$$
$$\overline{r} = \sum_i x_i \alpha_i + \qquad + \overline{r}_m \sum_i x_i \beta_i$$

$$r - \overline{r} = \sum_i x_i e_i + \left[\sum_i x_i \beta_i \right] \left(r_m - \overline{r}_m \right)$$

$$\left(r - \overline{r} \right)^2 = \sum_i \sum_j x_i x_j e_i e_j + 2 \left[\sum_i x_i \beta_i \right] \left(r_m - \overline{r}_m \right) \sum_i x_i e_i + \left(\sum_i x_i \beta_i \right)^2 \left(r_m - \overline{r}_m \right)^2$$

Aplicando o operador esperança:

$$\sigma_r^2 = \sum_i \sum_j x_i x_j E\left[e_i e_j\right] + \left(\sum_i x_i \beta_i\right)^2 E\left[\left(r_m - \overline{r}_m\right)^2\right]$$

$$\sigma_r^2 = \sum_i \sum_j x_i x_j \sigma_{io}^2 \delta_{ij} + \left(\sum_i x_i \beta_i\right)^2 \sigma_m^2$$

$$\sigma_r^2 = \sum_i x_i^2 \sigma_{io}^2 + \left(\sum_i x_i \beta_i\right)^2 \sigma_m^2$$

Onde $\sigma_{div} = \sqrt{\sum_i x_i^2 \sigma_{io}^2}$ é o risco diversificável e $\sigma_m = \sqrt{\left(\sum_i x_i \beta_i\right)^2 \sigma_m^2}$ é o risco de mercado. A diversificação não protege contra o risco de mercado, pois todos os ativos sobem e descem juntos. Mas ajuda para anular o risco diversificável. Suponha que utilizamos N ativos igualmente distribuídos, ou seja, $x_i = \frac{1}{N}$. Nesse caso $\sum_i x_i^2 \sigma_{io}^2 = \frac{1}{N^2} \sum_i \sigma_{io}^2 = \frac{1}{N}\left[\frac{1}{N}\sum_i \sigma_{io}^2\right] = \frac{1}{N}\overline{\sigma_o^2}$. Logo $\sigma_{div} = \frac{1}{\sqrt{N}}\sigma_{RMS}$ onde RMS = Root Mean Square implica em elevar ao quadrado, tirar a média e, finalmente, a raiz quadrada. Por outro lado $\sum_i x_i \beta_i = \frac{1}{N}\sum_i \beta_i = \overline{\beta}$, logo:

$$\sigma_r^2 = \frac{1}{N}\sigma_{RMS}^2 + \overline{\beta}^2 \sigma_m^2$$

Nesse caso se $N \to \infty$ o $\sigma_{div} \to 0$ mas $\sigma_r = \left|\overline{\beta}\right| \sigma_m$ se mantém não nulo. O risco diversificável cai com a diversificação mas cai com $\frac{1}{\sqrt{N}}$, ou seja, se aumentamos o número de ativos por um fator de 10 o risco cai apenas por um fator da ordem de 3.

Shrinkage:

Trata-se de mais um método que tenta limpar a matriz de covariância. Na nossa forma de ver estão usando o alvo errado, pois a matriz de covariância é mais robusta do que as previsões dos retornos futuros. Depois para limpar a matriz de covariância outros métodos baseados nos autovalores podem ser mais adequados. Nesse método se deseja dar mais peso aos elementos da diagonal, variância, da matriz de covariância do que aos fora da diagonal, covariância. Dessa forma se re-escreve a matriz de covariância como:

$$SV_{ij} = \lambda V_{ij} + \left(1-\lambda\right)\sigma_i^2\delta_{ij}$$

Nenhuma teoria a respeito de qual deve ser λ. A idéia é variar λ para trazer os valores de x do portfólio de Black o mais próximo de positivos possível. A implementação com Black é fácil – pois basta usar a SV_{ij} em lugar da V_{ij} e podemos ver imediatamente como a fronteira e o portfólio se move.

No caso particular de $\lambda = 0$ a matriz de covariância se torna diagonal.

Nesse caso $V^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_A^2} \end{pmatrix}$ e a matriz A será dada por:

$$A = \begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_n^2} \end{pmatrix} \begin{pmatrix} \bar{r}_1 & 1 \\ \bar{r}_2 & 1 \\ \vdots & \vdots \\ \bar{r}_A & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_A \\ 1 & 1 & \ldots & 1 \end{pmatrix} \begin{pmatrix} \frac{\bar{r}_1}{\sigma_1^2} & \frac{1}{\sigma_1^2} \\ \frac{\bar{r}_2}{\sigma_2^2} & \frac{1}{\sigma_2^2} \\ \vdots & \vdots \\ \frac{\bar{r}_A}{\sigma_A^2} & \frac{1}{\sigma_A^2} \end{pmatrix} = \begin{pmatrix} \sum_i \frac{\bar{r}_i^2}{\sigma_i^2} & \sum_i \frac{\bar{r}_i}{\sigma_i^2} \\ \sum_i \frac{\bar{r}_i}{\sigma_i^2} & \sum_i \frac{1}{\sigma_i^2} \end{pmatrix}$$

$$A = \begin{pmatrix} \sum_i \frac{\overline{r}_i^{\,2}}{\sigma_i^2} & \sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ \sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{1}{\sigma_i^2} \end{pmatrix}$$

$$A^{-1} = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i^{\,2}}{\sigma_i^2 \sigma_j^2} - \sum_i \sum_j \frac{\overline{r}_i\,\overline{r}_j}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \sum_i \frac{1}{\sigma_i^2} & -\sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ -\sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{\overline{r}_i^{\,2}}{\sigma_i^2} \end{pmatrix}$$

$$A^{-1} = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \sum_i \frac{1}{\sigma_i^2} & -\sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ -\sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{\overline{r}_i^{\,2}}{\sigma_i^2} \end{pmatrix}$$

$$\sigma^2 = \begin{pmatrix} \overline{r} & 1 \end{pmatrix} A^{-1} \begin{pmatrix} \overline{r} \\ 1 \end{pmatrix}$$

$$\sigma^2 = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \overline{r} & 1 \end{pmatrix} \begin{pmatrix} \sum_i \frac{1}{\sigma_i^2} & -\sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ -\sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{\overline{r}_i^{\,2}}{\sigma_i^2} \end{pmatrix} \begin{pmatrix} \overline{r} \\ 1 \end{pmatrix}$$

$$\sigma^2 = \frac{1}{\left[\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}\right]} \sum_i \frac{\left(\overline{r} - \overline{r}_i\right)^2}{\sigma_i^2}$$

$$\vec{X}_{Black} = M \begin{pmatrix} \overline{r} \\ 1 \end{pmatrix}, \text{ onde } M = V^{-1} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} A^{-1} .$$

$$M = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_A^2} \end{pmatrix} \begin{pmatrix} \overline{r}_1 & 1 \\ \overline{r}_2 & 1 \\ \vdots & \vdots \\ \overline{r}_A & 1 \end{pmatrix} \begin{pmatrix} \sum_i \frac{1}{\sigma_i^2} & -\sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ -\sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{\overline{r}_i^2}{\sigma_i^2} \end{pmatrix}$$

$$M = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \frac{\overline{r}_1}{\sigma_1^2} & \frac{1}{\sigma_1^2} \\ \frac{\overline{r}_2}{\sigma_2^2} & \frac{1}{\sigma_2^2} \\ \vdots & \vdots \\ \frac{\overline{r}_A}{\sigma_A^2} & \frac{1}{\sigma_A^2} \end{pmatrix} \begin{pmatrix} \sum_i \frac{1}{\sigma_i^2} & -\sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ -\sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{\overline{r}_i^2}{\sigma_i^2} \end{pmatrix}$$

$$M = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \frac{\overline{r}_1}{\sigma_1^2} & \frac{1}{\sigma_1^2} \\ \frac{\overline{r}_2}{\sigma_2^2} & \frac{1}{\sigma_2^2} \\ \vdots & \vdots \\ \frac{\overline{r}_A}{\sigma_A^2} & \frac{1}{\sigma_A^2} \end{pmatrix} \begin{pmatrix} \sum_i \frac{1}{\sigma_i^2} & -\sum_i \frac{\overline{r}_i}{\sigma_i^2} \\ -\sum_i \frac{\overline{r}_i}{\sigma_i^2} & \sum_i \frac{\overline{r}_i^2}{\sigma_i^2} \end{pmatrix}$$

$$\vec{X}_{Black} = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i\left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \frac{\overline{r}_1}{\sigma_1^2} & \frac{1}{\sigma_1^2} \\ \frac{\overline{r}_2}{\sigma_2^2} & \frac{1}{\sigma_2^2} \\ \vdots & \vdots \\ \frac{\overline{r}_A}{\sigma_A^2} & \frac{1}{\sigma_A^2} \end{pmatrix} \begin{pmatrix} \sum_i \frac{\left(\overline{r} - \overline{r}_i\right)}{\sigma_i^2} \\ -\sum_i \frac{\left(\overline{r} - \overline{r}_i\right)\overline{r}_i}{\sigma_i^2} \end{pmatrix}$$

$$x_k = \frac{\sum_j \frac{\left(\bar{r} - \bar{r}_j\right)\left(\bar{r}_k - \bar{r}_j\right)}{\sigma_k^2 \sigma_j^2}}{\sum_i \sum_j \frac{\bar{r}_i \left(\bar{r}_i - \bar{r}_j\right)}{\sigma_i^2 \sigma_j^2}}$$

Esse caso particular pode ser resolvido diretamente da seguinte forma:

Minimizar $\sigma^2 = \sum_i x_i^2 \sigma_i^2$ sujeito às restrições $\sum x_i = 1$, $\sum x_i \bar{r}_i = \bar{r}$.

O Lagrangeano do sistema é dado por:

$$\mathcal{L} = \sum_i x_i^2 \sigma_i^2 - \lambda_1 \left[\sum_i x_i \bar{r}_i - \bar{r} \right] - \lambda_2 \left[\sum_i x_i - 1 \right]$$

As condições de primeira ordem, $\frac{\partial \mathcal{L}}{\partial x_k} = 0$:

$$\frac{\partial}{\partial x_k} \mathcal{L} = 2 x_k \sigma_k^2 - \lambda_1 \bar{r}_k - \lambda_2 = 0$$

$$x_k = \begin{pmatrix} \frac{\bar{r}_k}{\sigma_k^2} & \frac{1}{\sigma_k^2} \end{pmatrix} \left[\frac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \right]$$

Aplicando as restrições obtemos o sistema de equações:

$$\sum_k x_k = \frac{1}{2}\left(\lambda_1 \sum_k \frac{\overline{r}_k}{\sigma_k^2} + \lambda_2 \sum_k \frac{1}{\sigma_k^2}\right) = 1$$

$$\sum_k x_k \overline{r}_k = \frac{1}{2}\left(\lambda_1 \sum_k \frac{\overline{r}_k^2}{\sigma_k^2} + \lambda_2 \sum_k \frac{\overline{r}_k}{\sigma_k^2}\right) = r_{\text{target}}$$

$$\begin{pmatrix} \sum_k \frac{\overline{r}_k}{\sigma_k^2} & \sum_k \frac{1}{\sigma_k^2} \\ \sum_k \frac{\overline{r}_k^2}{\sigma_k^2} & \sum_k \frac{\overline{r}_k}{\sigma_k^2} \end{pmatrix} \left[\frac{1}{2}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ r_{\text{target}} \end{pmatrix}$$

$$\frac{1}{2}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i \overline{r}_j}{\sigma_i^2 \sigma_j^2} - \sum_i \sum_j \frac{\overline{r}_i^2}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} \sum_j \frac{\overline{r}_j}{\sigma_j^2} & -\sum_j \frac{1}{\sigma_j^2} \\ -\sum_j \frac{\overline{r}_j^2}{\sigma_j^2} & \sum_j \frac{\overline{r}_j}{\sigma_j^2} \end{pmatrix} \begin{pmatrix} 1 \\ r_{\text{target}} \end{pmatrix}$$

$$\frac{1}{2}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \frac{1}{\sum_i \sum_j \frac{\overline{r}_i \left(\overline{r}_i - \overline{r}_j\right)}{\sigma_i^2 \sigma_j^2}} \begin{pmatrix} -\sum_j \frac{\overline{r}_j}{\sigma_j^2} & \sum_j \frac{1}{\sigma_j^2} \\ \sum_j \frac{\overline{r}_j^2}{\sigma_j^2} & -\sum_j \frac{\overline{r}_j}{\sigma_j^2} \end{pmatrix} \begin{pmatrix} 1 \\ r_{\text{target}} \end{pmatrix}$$

$$\frac{1}{2}\lambda_1 = \frac{\sum_j \frac{\left(\overline{r}_j - r_{\text{target}}\right)}{\sigma_j^2}}{\left(\sum_j \frac{\overline{r}_j}{\sigma_j^2}\right)^2 - \left(\sum_j \frac{\overline{r}_j^2}{\sigma_j^2}\right)\left(\sum_j \frac{1}{\sigma_j^2}\right)}$$

$$\frac{1}{2}\lambda_2 = \frac{\sum_j \frac{\overline{r}_j \left(r_{\text{target}} - \overline{r}_j\right)}{\sigma_j^2}}{\left(\sum_j \frac{\overline{r}_j}{\sigma_j^2}\right)^2 - \left(\sum_j \frac{\overline{r}_j^2}{\sigma_j^2}\right)\left(\sum_j \frac{1}{\sigma_j^2}\right)}$$

$$\frac{1}{2}\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}=\frac{1}{\sum_i\sum_j\frac{\overline{r}_i\left(\overline{r}_i-\overline{r}_j\right)}{\sigma_i^2\sigma_j^2}}\begin{pmatrix}-\sum_j\frac{\overline{r}_j}{\sigma_j^2} & \sum_j\frac{1}{\sigma_j^2}\\ \sum_j\frac{\overline{r}_j^{\,2}}{\sigma_j^2} & -\sum_j\frac{\overline{r}_j}{\sigma_j^2}\end{pmatrix}\begin{pmatrix}1\\ r_{\text{target}}\end{pmatrix}$$

$$x_k=\frac{1}{\left[\sum_i\sum_j\frac{\overline{r}_i\left(\overline{r}_i-\overline{r}_j\right)}{\sigma_i^2\sigma_j^2}\right]}\begin{pmatrix}\frac{\overline{r}_k}{\sigma_k^2} & \frac{1}{\sigma_k^2}\end{pmatrix}\begin{pmatrix}-\sum_j\frac{\overline{r}_j}{\sigma_j^2} & \sum_j\frac{1}{\sigma_j^2}\\ \sum_j\frac{\overline{r}_j^{\,2}}{\sigma_j^2} & -\sum_j\frac{\overline{r}_j}{\sigma_j^2}\end{pmatrix}\begin{pmatrix}1\\ r_{\text{target}}\end{pmatrix}$$

$$\vec{X}_{Black}=M\begin{pmatrix}\overline{r}\\ 1\end{pmatrix},\text{ onde } M=V^{-1}\begin{pmatrix}\overline{r}_1 & 1\\ \overline{r}_2 & 1\\ \vdots & \vdots\\ \overline{r}_A & 1\end{pmatrix}A^{-1}\ .$$

Portfólio de Máxima Diversificação [Maximização da Entropia]:

Sabemos que a entropia de Shannon é dada por $H=-\sum_i x_i\ln x_i$ com $\sum_i x_i=1$ e $x_i\geq 0$.

Podemos nos perguntar qual o portfólio que maximiza a entropia sem qualquer informação. Esse problema é expresso matematicamente como:

Maximizar $H=-\sum_i x_i\ln x_i$ sujeito às restrições $\sum_i x_i=1$ e $x_i\geq 0$.

Vamos esquecer a restrição $x_i\geq 0$ porque ela pode não COLAR,m i.e., mesmo sem ela os valores encontrados são positivos. Nesse caso o Lagrangeano é dado por:

$$\mathcal{L} = -\sum_i x_i \ln x_i - \lambda\left(\sum_i x_i - 1\right)$$

A solução será dada por

$$\frac{\partial}{\partial x_j}\mathcal{L} = -\sum_i \delta_{ij} \ln x_i - \sum_i x_i \frac{1}{x_i}\delta_{ij} - \lambda\sum_i \delta_{ij} = 0$$

$$-\ln x_j - 1 - \lambda = 0$$

$x_j = e^{-1-\lambda} \geq 0$ ou seja x_j é uma constante positiva, que obedece naturalmente à restrição $x_j \geq 0$.

O multiplicador de Lagrange é extraído de: $\sum_j x_j = e^{-1-\lambda}\sum_j 1 = Ne^{-1-\lambda} = 1$ logo $e^{-1-\lambda} = \frac{1}{N}$ e, portanto, $x_j = \frac{1}{N}$. Nesse caso a regra é: investir a mesma fração em todos os ativos. Embora $\begin{pmatrix} x_1 & x_2 & \cdots & x_N \end{pmatrix} = \begin{pmatrix} \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \end{pmatrix}$ seja considerado um portfólio de TOLOS, não é tão ruim assim porque é um portfólio que maximiza a diversificação, a entropia.

O próximo passo é exigir que se deseja obter um alvo de retorno, incluindo mais uma restrição dada por: $\sum_i r_i x_i = r_{\text{target}}$. O problema é:

Maximizar $H = -\sum_i x_i \ln x_i$ sujeito às restrições $\sum_i r_i x_i = r_{\text{target}}$, $\sum_i x_i = 1$ e $x_i \geq 0$.

Novamente, esquecendo a restrição $x_i \geq 0$, o Lagrangeano é dado por:

$$\mathcal{L} = -\sum_i x_i \ln x_i - \mu\left(\sum_i x_i - 1\right) - \lambda\left(\sum_i x_i \bar{r}_i - r_{\text{target}}\right)$$

A solução será dada por

$$\frac{\partial}{\partial x_j}\mathcal{L} = -\ln x_j - 1 - \mu - \lambda \bar{r}_j = 0$$

$$x_j = e^{-1-\mu} e^{-\lambda \bar{r}_j} = x_o e^{-\lambda \bar{r}_j}$$

Encontrando as constantes: x_o e λ: $\sum_j x_j = x_o \sum_j e^{-\lambda \bar{r}_j} = 1$ e $\sum_j x_j \bar{r}_j = x_o \sum_j \bar{r}_j e^{-\lambda \bar{r}_j} = r_{\text{target}}$. Nesse caso $x_o = \frac{1}{\sum_i e^{-\lambda \bar{r}_i}}$ e $\frac{\sum_j \bar{r}_j e^{-\lambda \bar{r}_j}}{\sum_i e^{-\lambda r_i}} = r_{\text{target}}$. Note que $e^{-\lambda \bar{r}_i} \geq 0$, logo $x_o \geq 0$ e $x_j = x_o e^{-\lambda \bar{r}_j} \geq 0$ e a restrição $x_j \geq 0$ está automaticamente satisfeita.

Para calcular λ: usamos a restrição $\frac{\sum_j \bar{r}_j e^{-\lambda \bar{r}_j}}{\sum_i e^{-\lambda r_i}} = r_{\text{target}}$ ou seja, $\sum_j \bar{r}_j e^{-\lambda \bar{r}_j} - \sum_i r_{\text{target}} e^{-\lambda r_i} = 0$ ou ainda $\sum_j \left(\bar{r}_j - r_{\text{target}} \right) e^{-\lambda \bar{r}_j} = 0$.

A estratégia para resolver o problema é:

1. Resolver para λ a equação $\sum_j \bar{r}_j e^{-\lambda \bar{r}_j} - \sum_i r_{\text{target}} e^{-\lambda r_i} = 0$ usando solver.
2. Determinado λ podemos calcular $x_o = \frac{1}{\sum_i e^{-\lambda \bar{r}_i}}$.
3. Finalmente calculamos cada um dos elementos $x_k = \frac{e^{-\lambda \bar{r}_k}}{\sum_i e^{-\lambda \bar{r}_i}}$.

Apêndice 1: Macro para encontrar a Fronteira Eficiente de Markowitz:

```
Sub Markowitz()
'
' Markowitz Macro
' Encontrar fronteira eficiente de Markowitz
'
' Keyboard Shortcut: Ctrl+Shift+M
'
' Estabelece as linhas em que o SOLVER será aplicado e o LOOP da MACRO:

For Row = 7 To 57

' Para limpar o SOLVER:

SolverReset

' Estabelecendo as duas restrições:

SolverAdd CellRef:=Cells(Row, "EK"), Relation:=2, FormulaText:="0"
SolverAdd CellRef:=Cells(Row, "EL"), Relation:=2, FormulaText:="0"

' Estabelecendo os parâmetros do SOLVER - especialmente a restrição de x ser positivo

SolverOptions MaxTime:=10000, Iterations:=10000, Precision:=0.00000001, AssumeLinear:=False, StepThru:=False, Estimates:=1, Derivatives:=1, SearchOption:=1, IntTolerance:=5, Scaling:=False, Convergence:=0.000001, AssumeNonNeg:=True

' Definindo quem é a célula alvo, que deve ser minimizada (2) através da variação de que células

SolverOk SetCell:=Cells(Row, "EM"), MaxMinVal:=2, ValueOf:="0", ByChange:=Range(Cells(Row, "BW"), Cells(Row, "EJ"))

' Os dois comandos a seguir são importantes para que o MACRO não fique perguntando em cada passo se a solução está boa

SolverFinish keepFinal:=1
SolverSolve userFinish:=True

' Para copiar os valores obtidos no passo anterior:

Range(Cells(Row, "BW"), Cells(Row, "EJ")).Select
Selection.AutoFill Destination:=Range(Cells(Row, "BW"), Cells(Row + 1, "EJ")), Type:=xlFillDefault

' Procedendo para linha seguinte:

Next Row

End Sub
```

Apêndice 2: Macro para encontrar a Fronteira Eficiente de Black. [Note que a fronteira de Black pode ser obtida analiticamente sem necessidade de uma Macro – mas também pode ser determinada pela Macro]

```
Sub Black()
'
' Black Macro
' Encontrar fronteira eficiente de Black
'
' Keyboard Shortcut: Ctrl+Shift+B
'
' Estabelece as linhas em que o SOLVER será aplicado e o LOOP da MACRO:

For Row = 7 To 57

' Para limpar o SOLVER:

SolverReset

' Estabelecendo as duas restrições:

SolverAdd CellRef:=Cells(Row, "EK"), Relation:=2, FormulaText:="0"
SolverAdd CellRef:=Cells(Row, "EL"), Relation:=2, FormulaText:="0"

' Estabelecendo os parâmetros do SOLVER – liberar a restrição de x ser positivo

SolverOptions MaxTime:=10000, Iterations:=10000, Precision:=0.00000001, AssumeLinear:=False, StepThru:=False, Estimates:=1, Derivatives:=1, SearchOption:=1, IntTolerance:=5, Scaling:=False, Convergence:=0.000001, AssumeNonNeg:=False

' Definindo quem é a célula alvo, que deve ser minimizada (2) através da variação de que células

SolverOk SetCell:=Cells(Row, "EM"), MaxMinVal:=2, ValueOf:="0", ByChange:=Range(Cells(Row, "BW"), Cells(Row, "EJ"))

' Os dois comandos a seguir são importantes para que o MACRO não fique perguntando em cada passo se a solução está boa

SolverFinish keepFinal:=1
SolverSolve userFinish:=True

' Para copiar os valores obtidos no passo anterior:

Range(Cells(Row, "BW"), Cells(Row, "EJ")).Select
Selection.AutoFill Destination:=Range(Cells(Row, "BW"), Cells(Row + 1, "EJ")), Type:=xlFillDefault

' Procedendo para linha seguinte:

Next Row

End Sub
```

Apêndice 3: Macro para encontrar a Fronteira Eficiente de Maximização da Entropia.

```
Sub Entropia()
'
' Entropia Macro
'
' Keyboard Shortcut: Ctrl+Shift+E
'

For Row = 9 To 57

' Para limpar o SOLVER:

SolverReset

' Definindo quem é a célula alvo, que deve ser zerada (2) através da variação de que células

SolverOk SetCell:=Cells(Row, "BV"), MaxMinVal:=3, ValueOf:="0", ByChange:=Cells(Row, "BW")

' Os dois comandos a seguir são importantes para que o MACRO não fique perguntando em cada passo se a solução está boa

SolverFinish keepFinal:=1
SolverSolve userFinish:=True

' Para copiar os valores obtidos no passo anterior:

Cells(Row, "BW").Select
Selection.AutoFill Destination:=Range(Cells(Row, "BW"), Cells(Row + 1, "BW")), Type:=xlFillDefault

' Procedendo para linha seguinte:

Next Row

End Sub
```

Apêndice xxx.

Otimização sem restrições:

Suponha que a função $f(\vec{x}):\mathbb{R}^n \to \mathbb{R}$ de classe $C^{k>2}$ tenha um extremo, máximo ou mínimo ou ponto de sela, em $\vec{x}=\vec{x}^*$. Vamos transformar o problema de n dimensões em um problema de uma dimensão através da parametrização: $\vec{x}=\vec{x}^*+t\vec{h}$ onde $t\in\mathbb{R}$ e $\vec{h}\neq\vec{0}$. Nesse caso a função de uma variável $g(t)=f\left(\vec{x}^*+t\vec{h}\right)$ tem extremo em $t=0$ para $\forall\,\vec{h}$.

A condição de primeira ordem para ser ponto de máximo, mínimo ou sela é que $g'(0)=0$. Derivando $g(t)$ obtemos:

$$\frac{d}{dt}g(t)=\frac{\partial f}{\partial x_1}\frac{d}{dt}\left(x_1^*+th_1\right)+\frac{\partial f}{\partial x_2}\frac{d}{dt}\left(x_2^*+th_2\right)+\cdots+\frac{\partial f}{\partial x_n}\frac{d}{dt}\left(x_n^*+th_n\right)$$

$$\frac{d}{dt}g(t)=h_1\frac{\partial f}{\partial x_1}+h_2\frac{\partial f}{\partial x_2}+\cdots+h_n\frac{\partial f}{\partial x_n}$$

Logo, para ser um ponto extremo é preciso que:

$$h_1\frac{\partial f}{\partial x_1}+h_2\frac{\partial f}{\partial x_2}+\cdots+h_n\frac{\partial f}{\partial x_n}=0$$

Agora, essa igualdade deve ser verdadeira para qualquer $\vec{h}$. Escolhendo $\vec{h}'=\begin{pmatrix}0 & 0 & \cdots & 0 & h_i & 0 & \cdots & 0\end{pmatrix}$ então $\sum_j h_j\frac{\partial f}{\partial x_j}=h_i\frac{\partial f}{\partial x_i}=0$ implica que $\frac{\partial}{\partial x_i}f\left(\vec{x}^*\right)=0$. Fazendo $i=1,2,\cdots,n$ percebe-se que $\frac{\partial}{\partial x_i}f\left(\vec{x}^*\right)=0$ para $\forall\,i$. Em termos vetoriais isso pode ser escrito da forma:

$$\nabla f\left(\vec{x}^*\right)=\vec{0}.$$

Condições de Segunda Ordem:

Se o ponto extremo dor de máximo então $g''(0)<0$, e se for de mínimo então $g''(0)>0$. Em outras palavras a condição de segunda ordem é dada pelo sinal da segunda derivada:

$sign\left[g''(0)\right] = -1$ é ponto de máximo, e $sign\left[g''(0)\right] = +1$ é ponto de mínimo. Calculando a segunda derivada temos:

$$\frac{d^2}{dt^2} g(t) = \frac{d}{dt} \sum_i h_i \frac{\partial f}{\partial x_i} = \sum_i h_i \frac{d}{dt} \frac{\partial f}{\partial x_i} = \sum_i h_i \sum_j h_j \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Ou seja:

$$\frac{d^2}{dt^2} g(t) = \sum_i \sum_j h_i \frac{\partial^2 f}{\partial x_j \partial x_i} h_j$$

A matriz Hessiana é definida como: $H_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}$. Trata-se de uma matriz simétrica pois $H_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = H_{ji}$. Podemos escrever essa derivada na forma de uma multiplicação de matrizes como $\frac{d^2}{dt^2} g(t) = \vec{h}' H \vec{h}$. Algumas pessoas usam a notação $H = \nabla^2 f$, mas não gostamos dessa notação na física porque é mesma do Laplaciano $\nabla^2 = \nabla \cdot \nabla = \sum_i \frac{\partial^2}{\partial x_i^2}$ que é um operador escalar e não tensorial como o Hessiano.

$$\sum_i \sum_j h_i H_{ij} h_j = \begin{pmatrix} h_1 & h_2 & \cdots & h_n \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}$$

Nesse caso a condição para o ponto extremo ser de máximo é que $\frac{d^2}{dt^2} g(t) = \vec{h}' H \vec{h} < 0$, o que significa que a matriz Hessiana deve ser definida negativa. Já para ser de mínimo a condição é que $\frac{d^2}{dt^2} g(t) = \vec{h}' H \vec{h} > 0$, o que significa que a matriz Hessiana deve ser definida positiva. Se a matriz Hessiana não for nem definida positiva, nem definida negativa, então o ponto é de sela.

Definição de matrizes simétricas:

Dizemos que a matriz $M_{n\times n}=\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{pmatrix}$ é definida positiva se $\vec{h}'M\vec{h}>0$ para qualquer $\vec{h}\neq\vec{0}$. A matriz será definida negativa se $\vec{h}'M\vec{h}<0$ para qualquer $\vec{h}\neq\vec{0}$.

Teorema 1:

Se M é simétrica e definida positiva ou negativa então todos os elementos da diagonal ou são positivos ou são negativos. Prova: como a propriedade é válida para qualquer $\vec{h}\neq\vec{0}$ basta escolher $\vec{h}=\begin{pmatrix} 0 & 0 & \cdots & 0 & h_k & 0 & \cdots & 0 \end{pmatrix}$ o que significa que $\vec{h}'M\vec{h}=M_{kk}h_k^2$. Logo se $\vec{h}'M\vec{h}>0$ então $M_{kk}>0$ e se $\vec{h}'M\vec{h}<0$ então $M_{kk}<0$ para qualquer $k\in[1,n]$.

Teorema 2: se M é definida positiva ou negativa então todas as submatrizes $M_k=\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1k} \\ m_{12} & m_{22} & \cdots & m_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ m_{1k} & m_{2k} & \cdots & m_{kk} \end{pmatrix}$ serão definidas positivas ou negativas.

Prova: como a propriedade é válida para qualquer $\vec{h}\neq\vec{0}$ basta escolher $\vec{h}=\begin{pmatrix} h_1 & h_2 & \cdots & h_k & 0 & 0 & \cdots & 0 \end{pmatrix}$ o que significa que $\vec{h}'M\vec{h}=\vec{h}_k'M_k\vec{h}_k$.

Teorema 3.

a. Se M é definida positiva então $\det M_k>0$, ou ainda, $sign[\det M_k]=1$ para qualquer $k\in[1,n]$.
b. Se M é definida negativa então $sign[\det M_k]=(-1)^{k+1}$, ou seja, o sinal do determinante vai alternando na forma $\begin{pmatrix} - & + & - & + & \cdots \end{pmatrix}$.

A melhor forma de prova esse teorema é diagonalizar as submatrizes, mas é instrutivo analisar o caso de uma matriz 2×2 e outra 3×3 com a técnica de completar quadrado.

Matriz 2×2:

$$q=\begin{pmatrix} h & k \end{pmatrix}\begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix}$$

$$q=\begin{pmatrix} h & k \end{pmatrix}\begin{pmatrix} m_{11}h+m_{12}k \\ m_{21}h+m_{22}k \end{pmatrix}=m_{11}h^2+2m_{12}hk+m_{22}k^2$$

$$q=m_{11}\left[h^2+2\frac{m_{12}}{m_{11}}hk\right]+m_{22}k^2$$

$$q=m_{11}\left[h^2+2\frac{m_{12}}{m_{11}}hk+\left(\frac{m_{12}}{m_{11}}\right)^2k^2-\left(\frac{m_{12}}{m_{11}}\right)^2k^2\right]+m_{22}k^2$$

$$q=m_{11}\left[h^2+2\frac{m_{12}}{m_{11}}hk+\left(\frac{m_{12}}{m_{11}}\right)^2k^2\right]+m_{22}k^2-\frac{m_{12}^2}{m_{11}}k^2$$

$$q=m_{11}\left(h+\frac{m_{12}}{m_{11}}k\right)^2+\left[\frac{m_{11}m_{22}-m_{12}^2}{m_{11}}\right]k^2$$

$q=\left(\det M_1\right)h^{*2}+\dfrac{\left(\det M_2\right)}{\left(\det M_1\right)}k^2$ onde $h^*=h+\dfrac{m_{12}}{m_{11}}k$.

Se $q>0$ e como $h^{*2}>0$ e $k^2>0$ é preciso que $\det M_1>0$ e $\dfrac{\det M_2}{\det M_1}>0$ o que implica em $\det M_1>0$ e $\det M_2>0$. Por outro se $q<0$ é preciso que $\det M_1<0$ e $\dfrac{\det M_2}{\det M_1}<0$ o que implica em $\det M_1<0$ e $\det M_2>0$.

Matriz 3×3:

$$q=\begin{pmatrix}h & k & l\end{pmatrix}\begin{pmatrix}m_{11} & m_{12} & m_{13}\\ m_{12} & m_{22} & m_{23}\\ m_{13} & m_{23} & m_{33}\end{pmatrix}\begin{pmatrix}h\\ k\\ l\end{pmatrix}$$

$$q=\begin{pmatrix}h & k & l\end{pmatrix}\begin{pmatrix}m_{11}h+m_{12}k+m_{13}l\\ m_{12}h+m_{22}k+m_{23}l\\ m_{13}h+m_{23}k+m_{33}l\end{pmatrix}$$

$$q=m_{11}h^2+2m_{12}hk+2m_{13}hl+m_{22}k^2+2m_{23}kl+m_{33}l^2$$

$$q=m_{11}h^2+2\left[m_{12}k+m_{13}l\right]h+m_{22}k^2+2m_{23}kl+m_{33}l^2$$

$$q=m_{11}\left[h^2+2\frac{1}{m_{11}}\left(m_{12}k+m_{13}l\right)h+\frac{1}{m_{11}^2}\left(m_{12}k+m_{13}l\right)^2\right]+$$

$$-\frac{1}{m_{11}}\left(m_{12}k+m_{13}l\right)^2+m_{22}k^2+2m_{23}kl+m_{33}l^2$$

$$q=m_{11}\left(h+\frac{m_{12}k+m_{13}l}{m_{11}}\right)^2+m_{22}k^2-\frac{1}{m_{11}}\left(m_{12}k+m_{13}l\right)^2+2m_{23}kl+m_{33}l^2$$

$$q=m_{11}\left(h+\frac{m_{12}k+m_{13}l}{m_{11}}\right)^2+$$

$$+\frac{1}{m_{11}}\left(m_{11}m_{22}-m_{12}^2\right)k^2+2\frac{1}{m_{11}}\left(m_{11}m_{23}-m_{12}m_{13}\right)kl+\frac{1}{m_{11}}\left(m_{11}m_{33}-m_{13}^2\right)l^2$$

$$q = m_{11}\left(h + \frac{m_{12}k + m_{13}l}{m_{11}}\right)^2 +$$
$$+\frac{\left(m_{11}m_{22} - m_{12}^2\right)}{m_{11}}\left[k^2 + 2\frac{\left(m_{11}m_{23} - m_{12}m_{13}\right)}{\left(m_{11}m_{22} - m_{12}^2\right)}kl + \frac{\left(m_{11}m_{23} - m_{12}m_{13}\right)^2}{\left(m_{11}m_{22} - m_{12}^2\right)^2}l^2\right] +$$
$$-\frac{\left(m_{11}m_{22} - m_{12}^2\right)}{m_{11}}\frac{\left(m_{11}m_{23} - m_{12}m_{13}\right)^2}{\left(m_{11}m_{22} - m_{12}^2\right)^2}l^2 + \frac{1}{m_{11}}\left(m_{11}m_{33} - m_{13}^2\right)l^2$$

$$q = m_{11}\left(h + \frac{m_{12}k + m_{13}l}{m_{11}}\right)^2 + \frac{\left(m_{11}m_{22} - m_{12}^2\right)}{m_{11}}\left[k + \frac{\left(m_{11}m_{23} - m_{12}m_{13}\right)l}{\left(m_{11}m_{22} - m_{12}^2\right)}\right]^2 +$$
$$+\frac{1}{m_{11}\left(m_{11}m_{22} - m_{12}^2\right)}\left[\left(m_{11}m_{22} - m_{12}^2\right)\left(m_{11}m_{33} - m_{13}^2\right) - \left(m_{11}m_{23} - m_{12}m_{13}\right)^2\right]l^2$$

$$q = m_{11}\left(h + \frac{m_{12}k + m_{13}l}{m_{11}}\right)^2 + \frac{\left(m_{11}m_{22} - m_{12}^2\right)}{m_{11}}\left[k + \frac{\left(m_{11}m_{23} - m_{12}m_{13}\right)l}{\left(m_{11}m_{22} - m_{12}^2\right)}\right]^2 +$$
$$+\frac{1}{m_{11}\left(m_{11}m_{22} - m_{12}^2\right)}\left[\begin{array}{l} m_{11}^2m_{22}m_{33} - m_{11}m_{22}m_{13}^2 - m_{11}m_{12}^2m_{33} + m_{12}^2m_{13}^2 \\ -m_{11}^2m_{23}^2 + 2m_{11}m_{12}m_{13}m_{23} - m_{12}^2m_{13}^2 \end{array}\right]l^2$$

$$q = m_{11}\left(h + \frac{m_{12}k + m_{13}l}{m_{11}}\right)^2 + \frac{\left(m_{11}m_{22} - m_{12}^2\right)}{m_{11}}\left[k + \frac{\left(m_{11}m_{23} - m_{12}m_{13}\right)l}{\left(m_{11}m_{22} - m_{12}^2\right)}\right]^2 +$$
$$+\frac{\left(m_{11}m_{22}m_{33} + 2m_{12}m_{13}m_{23} - m_{11}m_{23}^2 - m_{22}m_{13}^2 - m_{12}^2m_{33}\right)}{\left(m_{11}m_{22} - m_{12}^2\right)}l^2$$

$$q = (\det M_1) h^{*2} + \frac{(\det M_2)}{(\det M_1)} k^{*2} + \frac{(\det M_3)}{(\det M_2)} l^2$$

Então se $q > 0$ exigimos que $\det M_1 > 0$, $\frac{\det M_2}{\det M_1} > 0$ e $\frac{\det M_3}{\det M_2} > 0$ o que implica em $\det M_1 > 0$, $\det M_2 > 0$ e $\det M_3 > 0$. Mas se $q < 0$ é preciso que $\det M_1 < 0$, $\frac{\det M_2}{\det M_1} < 0$ e $\frac{\det M_3}{\det M_2} < 0$ o que implica em $\det M_1 < 0$, $\det M_2 > 0$ e $\det M_3 < 0$.

No entanto a melhor forma de provar o caso geral é usar a técnica de diagonalização de matrizes.

Já sabemos que se a matriz M é definida positiva ou negativa então as sub-matrizes M_k possuem a mesma definição. Sabemos que uma transformação de similaridade do tipo $M' = S^{-1} M \, S$ preserva o determinante e queremos descobrir o sinal de $q = \vec{h}_k' M_k \vec{h}_k$. Seja S_k a matriz que diagonaliza M_k, simétrica, ou seja $S_k' M_k S_k = D_k$, então podemos fazer $q = \vec{h}_k' S_k S_k' M_k S_k S_k' \vec{h}_k$, ou seja, $q = \vec{h}_k' S_k D_k S_k' \vec{h}_k$. Agora definimos $\vec{h}_k^* = S_k' \vec{h}_k$ logo $\overline{\vec{h}_k^*} = \vec{h}_k' S_k$ logo $q = \overline{\vec{h}_k^*} D_k \vec{h}_k^*$.

Mas $D_k = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{pmatrix}$ então $q = \sum_{i=1}^{k} \lambda_i h_i^{*2}$. Logo se $q > 0$ então $\lambda_i > 0 \quad \forall i$, mas se se $q < 0$ então $\lambda_i < 0 \quad \forall i$. Se $\lambda_i > 0 \quad \forall i$ então $\det D_k > 0$, mas se $\lambda_i < 0 \quad \forall i$ então $sign[\det D_k] = (-1)^k$. Como a transformação de similaridade preserva os determinantes então, a condição $q > 0$ implica em $sign[\det M_k] = +1 \quad \forall k$ e a condição $q < 0$ implica em $sign[\det M_k] = (-1)^2 \quad \forall k$.

Note que a a definição da matriz Hessiana também define a concavidade da função. Se $f\left(\vec{x}\right)$ é estritamente côncava ou convexa então:

$$\begin{aligned} f\left[\left(1-\lambda\right)\vec{x}+\lambda\vec{x}'\right] &> \left(1-\lambda\right) f\left(\vec{x}\right)+\lambda f\left(\vec{x}'\right) \quad concava \\ f\left[\left(1-\lambda\right)\vec{x}+\lambda\vec{x}'\right] &< \left(1-\lambda\right) f\left(\vec{x}\right)+\lambda f\left(\vec{x}'\right) \quad convexa \end{aligned}$$

para $\forall\,\vec{x}\ e\ \vec{x}'$. Então também vale para $\vec{x}'=\vec{x}+\vec{h}$ com $\vec{h}\neq\vec{0}$ mas com $\left\|\vec{h}\right\|\to 0$. Nesse caso $\left(1-\lambda\right)\vec{x}+\lambda\vec{x}'=\left(1-\lambda\right)\vec{x}+\lambda\vec{x}+\lambda\vec{h}=\vec{x}+\lambda\vec{h}$ então:

$$\begin{aligned} f\left[\vec{x}+\lambda\vec{h}\right] &> \left(1-\lambda\right) f\left(\vec{x}\right)+\lambda f\left(\vec{x}+\vec{h}\right) \quad concava \\ f\left[\vec{x}+\lambda\vec{h}\right] &< \left(1-\lambda\right) f\left(\vec{x}\right)+\lambda f\left(\vec{x}+\vec{h}\right) \quad convexa \end{aligned}$$

Agora, por série de Taylor [ver apêndice convergência e série de Taylor] até segunda ordem sabemos que:

$$f\left(\vec{x}+\vec{h}\right)= f\left(\vec{x}\right)+\left(\vec{h}\cdot\nabla\right) f\left(\vec{x}\right)+\frac{1}{2}\left(\vec{h}\cdot\nabla\right)^2 f\left(\vec{x}\right)$$

$$f\left(\vec{x}+\vec{h}\right)= f\left(\vec{x}\right)+\sum_i h_i\partial_i f\left(\vec{x}\right)+\frac{1}{2}\sum_i\sum_j h_i h_j\partial_{ji} f\left(\vec{x}\right)$$

$$f\left(\vec{x}+\vec{h}\right)= f\left(\vec{x}\right)+\vec{h}'\nabla f\left(\vec{x}\right)+\frac{1}{2}\vec{h}'H\vec{h}$$

Fazendo $\vec{h}\to\lambda\vec{h}$ temos também que:

$$f\left(\vec{x}+\lambda\vec{h}\right)= f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda^2}{2}\vec{h}'H\vec{h}$$

Logo:

$$\begin{aligned} f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda^2}{2}\vec{h}'H\vec{h} &> \left(1-\lambda\right) f\left(\vec{x}\right)+\lambda\left[f\left(\vec{x}\right)+\vec{h}'\nabla f\left(\vec{x}\right)+\frac{1}{2}\vec{h}'H\vec{h}\right] \quad concava \\ f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda^2}{2}\vec{h}'H\vec{h} &< \left(1-\lambda\right) f\left(\vec{x}\right)+\lambda\left[f\left(\vec{x}\right)+\vec{h}'\nabla f\left(\vec{x}\right)+\frac{1}{2}\vec{h}'H\vec{h}\right] \quad convexa \end{aligned}$$

$$f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda^2}{2}\vec{h}'H\vec{h} > f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda}{2}\vec{h}'H\vec{h} \quad concava$$
$$f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda^2}{2}\vec{h}'H\vec{h} < f\left(\vec{x}\right)+\lambda\vec{h}'\nabla f\left(\vec{x}\right)+\frac{\lambda}{2}\vec{h}'H\vec{h} \quad convexa$$

$$\frac{\lambda^2}{2}\vec{h}'H\vec{h} > \frac{\lambda}{2}\vec{h}'H\vec{h} \quad concava$$
$$\frac{\lambda^2}{2}\vec{h}'H\vec{h} < \frac{\lambda}{2}\vec{h}'H\vec{h} \quad convexa$$

$$\frac{\lambda(1-\lambda)}{2}\vec{h}'H\vec{h} < 0 \quad concava$$
$$\frac{\lambda(1-\lambda)}{2}\vec{h}'H\vec{h} > 0 \quad convexa$$

Como $\lambda \in (0,1)$ então $\lambda(1-\lambda) > 0$ e a concavidade é definida por:

$$\vec{h}'H\vec{h} < 0 \quad concava$$
$$\vec{h}'H\vec{h} > 0 \quad convexa$$

ou seja, se o Hessiano for definido negativo a função é côncava e se for definido positivo a função é convexa. É intuitivo que funções côncavas possuam máximos enquanto funções convexas possuam mínimo.

Otimização com restrição de igualdade:

Suponha agora a seguinte problema: otimizar $f\left(\vec{x}\right)$ sujeito à restrição $g\left(\vec{x}\right)=c$, ou $g\left(\vec{x}\right)-c=0$. Agora nem todo $\vec{x}$ é permitido, só os que satisfazem à restrição $g\left(\vec{x}\right)=c$.

Se $\frac{\partial g}{\partial x_n}\left(\vec{x}^*\right)\neq 0$ então existe uma função inversa local implícita tal que $x_n=\varphi\left(x_1,x_2,\cdots,x_{n-1}\right)$ que pode ser utilizada para eliminar a variável x_n e caímos em funções de $n-1$ variáveis dadas por:

$$G\left(x_1,x_2,\cdots,x_{n-1}\right)=g\left[x_1,x_2,\cdots,x_{n-1},\varphi\left(x_1,x_2,\cdots,x_{n-1}\right)\right]$$

e

$$F\left(x_1,x_2,\cdots,x_{n-1}\right)=f\left[x_1,x_2,\cdots,x_{n-1},\varphi\left(x_1,x_2,\cdots,x_{n-1}\right)\right]$$

Note que a F já inclui a restrição, automaticamente satisfeita quando fizemos $x_n=\varphi\left(x_1,x_2,\cdots,x_{n-1}\right)$. Então as condições de ponto extremo da F são as usuais:

$$\frac{\partial}{\partial x_i}F=0 \quad i\in\left[1,2,\cdots,n-1\right] \text{ e } \begin{array}{l} \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\frac{\partial^2 F}{\partial x_i\partial x_j}h_i h_j<0 \quad \max \\ \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\frac{\partial^2 F}{\partial x_i\partial x_j}h_i h_j>0 \quad \min \end{array}$$

Agora $\frac{\partial G}{\partial x_i}=\frac{\partial g}{\partial x_i}+\frac{\partial g}{\partial x_n}\frac{\partial \varphi}{\partial x_i}=0$ pois $G=c$, o que significa que $\frac{\partial \varphi}{\partial x_i}=-\frac{\left(\frac{\partial g}{\partial x_i}\right)}{\left(\frac{\partial g}{\partial x_n}\right)}$. Por outro lado

$\frac{\partial F}{\partial x_i}=\frac{\partial f}{\partial x_i}+\frac{\partial f}{\partial x_n}\frac{\partial \varphi}{\partial x_i}=0$, então $\frac{\partial f}{\partial x_i}-\frac{\left(\frac{\partial f}{\partial x_n}\right)}{\left(\frac{\partial g}{\partial x_n}\right)}\frac{\partial g}{\partial x_i}=0$. Como $\frac{\left(\frac{\partial f}{\partial x_n}\right)}{\left(\frac{\partial g}{\partial x_n}\right)}=\lambda$ é uma constante, calculada no ponto $\vec{x}^*$ temos que:

$\frac{\partial f}{\partial x_i}-\lambda\frac{\partial g}{\partial x_i}=0$ ou ainda que $\frac{\partial}{\partial x_i}\left[f-\lambda g\right]=0 \quad i=1,2,\cdots,n-1$.

Por outro lado $\frac{\partial}{\partial x_n}\left[f-\lambda g\right]=\frac{\partial f}{\partial x_n}-\lambda\frac{\partial g}{\partial x_n}=\frac{\partial f}{\partial x_n}-\frac{\left(\frac{\partial f}{\partial x_n}\right)}{\left(\frac{\partial g}{\partial x_n}\right)}\frac{\partial g}{\partial x_n}=\frac{\partial f}{\partial x_n}-\frac{\partial f}{\partial x_n}=0$ automaticamente. Portanto podemos estender a condição para $\frac{\partial}{\partial x_i}\left[f-\lambda\left(g-c\right)\right]=0 \quad i=1,2,\cdots,n$. Definindo a função Lagrangeana dada por:

$\mathcal{L}(\lambda,\vec{x}) = f - \lambda\left[g(\vec{x}) - c\right]$ vemos que as condições de primeira ordem para um extremo são que $\frac{\partial}{\partial x_i}\mathcal{L}(\lambda,\vec{x}) = 0 \quad i = 1,2,\cdots,n$. Essa equação deve ser resolvida em conjunto com a restrição, mas notamos que $\frac{\partial}{\partial \lambda}\left[f - \lambda(g-c)\right] = 0$ implica em $(g-c) = 0$, ou seja, a própria restrição. Usando a função Lagrangeana o nosso problema é resolver o conjunto de equações simultâneas:

$$\frac{\partial}{\partial x_i}\mathcal{L}(\lambda,\vec{x}) = 0 \quad i = 1,2,\cdots,n$$

$$\frac{\partial}{\partial \lambda}\mathcal{L}(\lambda,\vec{x}) = 0$$

A condição de segunda ordem é um pouco mais complicada. Vamos usar a notação $\frac{\partial f}{\partial x_j} = f_j$ para simplificar. Nesse caso temos:

$$F_j = f_j + f_n\varphi_j$$
$$G_j = g_j + g_n\varphi_j$$

Sabemos que $G_j = g_j + g_n\varphi_j = 0$, logo $F_j = F_j - \lambda G_j$ e $F_{ij} = F_{ij} - \lambda G_{ij}$. Por outro lado:

$$F_{ij} = \left(f_{ij} + f_{nj}\varphi_i\right) + \left(f_{in} + f_{nn}\varphi_i\right)\varphi_j + f_n\varphi_{ij}$$
$$\lambda G_{ij} = \lambda\left(g_{ij} + g_{nj}\varphi_i\right) + \lambda\left(g_{in} + g_{nn}\varphi_i\right)\varphi_j + \lambda g_n\varphi_{ij}$$
$$F_{ij} - \lambda G_{ij} = \left(f_{ij} - \lambda g_{ij}\right) + \left(f_{in} - \lambda g_{in}\right)\varphi_j + \left(f_{nj} - \lambda g_{nj}\right)\varphi_i + \left(f_{nn} - \lambda g_{nn}\right)\varphi_i\varphi_j + \left(f_n - \lambda g_n\right)\varphi_{ij}$$

Agora:

$$f_n - \lambda g_n = f_n - \frac{f_n}{g_n}g_n = 0$$

Então:

$$\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(F_{ij}-\lambda G_{ij}\right)h_ih_j=\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(f_{ij}-\lambda g_{ij}\right)h_ih_j+\sum_{i=1}^{n-1}\left(f_{in}-\lambda g_{in}\right)h_i\sum_{j=1}^{n-1}\varphi_jh_j+$$
$$+\sum_{j=1}^{n-1}\left(f_{nj}-\lambda g_{nj}\right)h_j\sum_{i=1}^{n-1}\varphi_ih_i+\left(f_{nn}-\lambda g_{nn}\right)\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\varphi_ih_i\varphi_jh_j$$

Agora definindo um $h_n=\sum_{i=1}^{n-1}\varphi_ih_i$ e usando o fato de que $\varphi_i=-\frac{g_i}{g_n}$ temos que $h_n=-\frac{1}{g_n}\sum_{i=1}^{n-1}g_ih_i$, ou seja, $\sum_{i=1}^{n-1}g_ih_i+g_nh_n=0$, i.e., $\sum_{i=1}^{n}g_ih_i=0$ no garante que $\vec{h}$ está na curva $g\left(\vec{x}+\vec{h}\right)=c$. Nesse caso temos que:

$$\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(F_{ij}-\lambda G_{ij}\right)h_ih_j=\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(f_{ij}-\lambda g_{ij}\right)h_ih_j+\sum_{i=1}^{n-1}\left(f_{in}-\lambda g_{in}\right)h_ih_n+$$
$$+\sum_{j=1}^{n-1}\left(f_{nj}-\lambda g_{nj}\right)h_jh_n+\left(f_{nn}-\lambda g_{nn}\right)h_nh_n$$

Assim concluímos que:

$$\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}F_{ij}h_ih_j=\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(F_{ij}-\lambda G_{ij}\right)h_ih_j=\sum_{i=1}^{n}\sum_{j=1}^{n}\left(f_{ij}-\lambda g_{ij}\right)h_ih_j$$

Definição de matrizes simétricas com restrição.

Suponha agora que os vetores possíveis devem obedecer a uma restrição do tipo: $f_1h+f_2k=0$. Nesse caso as duas variáveis h e k não são mais independentes, mas devem obedecer a relação $h=-\frac{f_2}{f_1}k$. Assim $q=A_{11}h^2+2A_{12}hk+A_{22}k^2=A_{11}\frac{f_2^2}{f_1^2}k^2-2A_{12}\frac{f_2}{f_1}k^2+A_{22}k^2$

$$q=\left[f_2^2A_{11}-2f_1f_2A_{12}+f_1^2A_{22}\right]\frac{k^2}{f_1^2}$$

E o sinal de q será definido pelo sinal de $sign(q) = sign\left(f_2^2 A_{11} - 2 f_1 f_2 A_{12} + f_1^2 A_{22} \right)$. Agora percebemos que $\det\begin{pmatrix} 0 & f_1 & f_2 \\ f_1 & A_{11} & A_{12} \\ f_2 & A_{12} & A_{22} \end{pmatrix} = -\left[f_2^2 A_{11} - 2 f_1 f_2 A_{12} + f_1^2 A_{22} \right]$ logo em lugar de encontrar os sinais de dois subdeterminantes encontramos o sinal de apenas um mas com uma orla.

Generalizando:

Sem restrição com uma matriz Hessiana $n \times n$ da dorma $H = \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{12} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{1n} & H_{2n} & \cdots & H_{nn} \end{pmatrix}$ então a matriz é definida positiva se $\det H_1 = \det(H_{11})$, $\det H_2 = \det\begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix}$, $\det H_k = \det\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1k} \\ H_{12} & H_{22} & \cdots & H_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ H_{1k} & H_{2k} & \cdots & H_{kk} \end{pmatrix}$ até $\det H_n = \det\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{12} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ H_{1n} & H_{2n} & \cdots & H_{nn} \end{pmatrix}$ forem todos positivos. A matriz será definida negativa se $\det H_{2k+1} < 0$ e $\det H_{2k} > 0$.

Com restrição: $\tilde{H} = \begin{pmatrix} 0 & f_1 & \cdots & f_n \\ f_1 & H_{11} & \cdots & H_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ f_n & H_{1n} & \cdots & H_{nn} \end{pmatrix}$ será definido positivo se os determinantes $\det \tilde{H}_3 < 0$, $\det \tilde{H}_4 < 0$, ..., $\det \tilde{H}_{n+1} < 0$

Note que o Hessiano orlado poderia sair diretamente do Lagrangeano se incluirmos λ como uma variável pois:

$$\frac{\partial^2 \boldsymbol{\mathcal{L}}\left(\lambda,\vec{x}\right)}{\partial x_i \partial \lambda} = -\frac{\partial g\left(\vec{x}\right)}{\partial x_i} \text{ e } \frac{\partial^2 \boldsymbol{\mathcal{L}}\left(\lambda,\vec{x}\right)}{\partial \lambda^2} = 0$$