UNIVERSITAT DE BARCELONA

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

SELECCIÓN DE CARTERA ÓPTIMA COMO APLICACIÓN DE LAS CONDICIONES KARUSH-KUHN-TUKER

Autor: Yihao Yu

Director: Dr. José Manuel Concuera Valverde

Realitzat a: Departamente de Matemàtiques i Informàtica

Barcelona, 20 de junio de 2019

Abstract[1]

The Karush-Kuhn-Tucker conditions (in short, the KKT conditions), an extension of the well-known Lagrange multipliers method, have been developed to solve optimization problems in a more general sense, that is, including both inequalities and constraints. On the other hand, the selection of an optimal portfolio conforming the requirements of each investor, requesting a maximum return, a minimum risk or a balance between these two aspects, can be solved with the application of the KKT conditions.

Resumen

Las condiciones Karush-Kuhn-Tucker (en breve, las condiciones KKT), una extensión del conocido método de multiplicadores de Lagrange, han sido desarrolladas para resolver los problemas de optimización en un sentido más general, es decir, incluyendo tanto desigualdades como restricciones. Por otro lado, la selección de una cartera óptima conformando los requisitos de cada inversor, pidiendo una rentabilidad máxima, un riesgo mínimo o un equilibrio entre esos dos aspectos, puede ser resuelta con la aplicación de las condiciones KKT.

[1] 2010 Mathematics Subject Classification. 37N40, 46N10, 93E20

Agradecimiento

Primero de todo, quiero agradecer a mi tutor, José Manuel Concuera Valverde, su apoyo, dedicación y correcciones del trabajo. También quisiera agradecer al profesor Ricardo García López, por sus conocimientos en el campo de la Geometría diferencial y sus recomendaciones sobre el uso de ciertas herramientas. Sin su apoyo y el de mi familia, la redacción de este trabajo sería imposible.

Índice

Capítulo 1. Introducción

El multiplicador es una herramienta potente para enfrentarse a los problemas de optimización. Esta técnica permite encontrar el máximo o el mínimo de una función objetiva, que normalmente es multivariable y que sigue la forma siguiente:

$$f(x), donde\ x \in \mathbb{R}^n$$

Cuando hay alguna o varias restricciones en los valores de las variables independientes que se pueden usar, estas pueden ser las igualdades, las desigualdades o, en el caso más general, una combinación de ambas.

Por lo tanto, este artículo pretende, en primer lugar, abordar una introducción sobre qué es un multiplicador de Lagrange con las restricciones de igualdades y, como el caso más simple, para esta introducción, se introducirá la demostración de esta técnica, tanto a nivel geométrico como a nivel analítico. Después se aplicará esta idea a los casos más generales, en los que las restricciones serán una combinación de igualdades y desigualdades.

En segundo lugar, este artículo, presentará el famoso teorema del portafolio, publicada por el Dr. Harry Markowitz en la revista "The Journal of Finance" en el marzo del año 1952, a partir del cual se pretende aplicar el multiplicador de Lagrange como una técnica para resolver el problema de optimización de portafolio que consta de tres partes: la minimización de la varianza con una esperanza fijada, la maximización de la esperanza con una varianza conocida y la optimización del *trade-off* de esperanza y varianza a la vez.

Estructura del trabajo

Este trabajo se dividirá en dos capítulos principales: el capítulo 2, que se enfocará en la introducción del multiplicador de Lagrange, su demostración y su extensión a las restricciones de desigualdades (las condiciones KKT); y el capítulo 3 que pondrá más énfasis en el teorema de portafolio de Markowitz y lo tomará como caso concreto del problema de optimización. Se usará la técnica del multiplicador de Lagrange y su extensión, las condiciones KKT, para resolver este problema.

Capítulo 2. Multiplicador de Lagrange

En la optimización matemática, el método de los multiplicadores de Lagrange (o el método de los multiplicadores indeterminados de Lagrange, llamado así por Joseph-Louis Lagrange) es una estrategia para encontrar los máximos y mínimos locales de una función sujeta a restricciones de igualdad (es decir, sujeto a la condición de que una o más ecuaciones deben ser satisfechas exactamente por los valores elegidos de las variables). La gran ventaja de este método es que permite que la optimización se resuelva sin parametrización explícita en términos de las restricciones. Como resultado, el método de los multiplicadores de Lagrange se usa ampliamente para resolver problemas complejos de optimización restringidos.

El método se puede resumir de la siguiente manera:

1. Aislar cualquier posible punto singular del conjunto de soluciones de las ecuaciones restrictivas

2. Encontrar todos los puntos estacionarios de la función de Lagrange

3. Establecer cuáles de esos puntos estacionarios y puntos singulares son máximos globales (o mínimos, en caso de problemas de minimización) de la función objetivo.

2.1 Introducción

Para comenzar, estudiaremos el caso más simple: minimización con restricciones de igualdades solamente. En concreto, nos enfocaremos en el problema de la siguiente forma:

$$\text{Minimizar } f(x)$$

$$\text{Sujeto a} \quad h(x) = 0;$$

Donde $x \in \Omega \subset R^n$, las funciones$f, h = (h_1, h_2 .....h_m)$ son funciones continuas y, además, suponemos que sus primeras y segundas derivadas también lo son, por lo que son funciones de tipo C^2. Cuando $f,\ h_i$ son funciones lineales, el problema se convierte en una programación lineal, que podemos resolver con facilidad. por lo que aquí nos centramos en el caso contrario, en el que todas esas funciones no sean lineales.

Para poder simplificar el problema consideraremos n=2 y m=1, de modo que el problema se convierte en:

$$\text{Minimizar } f(x, y)$$

$$\text{Sujeto a} \quad h(x, y) = 0; \text{ Donde } x, y \in R$$

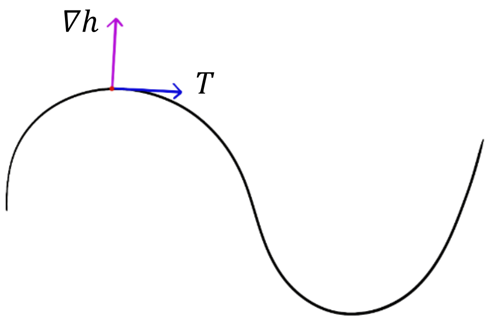
La restricción de igualdad $h(x, y)$ define una curva. Si diferenciamos esta función respecto a la variable x, tenemos:

$$\frac{\partial h}{\partial x} + \frac{\partial h}{\partial y}\frac{dy}{dx} = 0$$

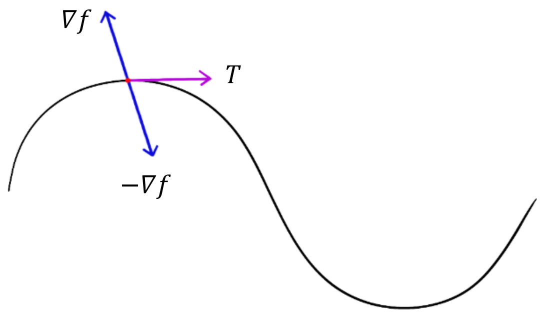
La recta tangente de la curva es $T(x, y) = \left(1, \frac{dy}{dx}\right)$ y el gradiente de la curva es $\nabla h = \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}\right)$. Así, la ecuación de arriba afirma que:

$$T \cdot \nabla h = 0$$

Evidentemente, el tangente siempre tiene que ser ortogonal al gradiente. Suponiendo que estamos en un punto de la curva, cualquier movimiento de los puntos debe estar en la dirección de la recta tangente que pase por el mismo punto.



Para incrementar o disminuir el valor de $f(x,y)$, los movimientos en la curva deben tener un componente en la dirección de la gradiente de f, ya que el gradiente indica en qué dirección el campo podría variar su valor a una mayor velocidad, que quiere decir, $\nabla f \cdot T \neq 0$.



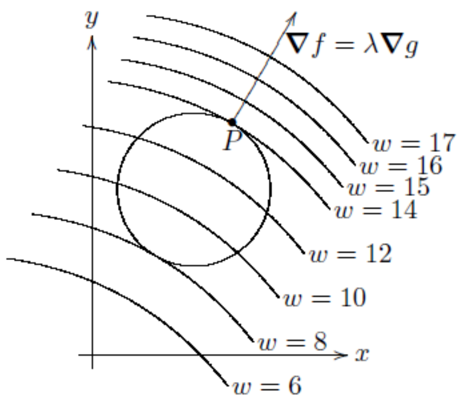
En un extremo de la función f, un movimiento diferencial no debe tener el componente en la dirección de ∇f ya que, en este punto, f goza su valor máximo. Entonces T es ortogonal a ∇f o, dicho en otras palabras, cumple la condición:

$$\nabla f \cdot T = 0$$

Ahora, ya hemos visto que T es ortogonal a los dos gradientes, ∇f y ∇h en un punto extremo. Eso significa que ∇f y ∇h deben estar paralelos. Por lo tanto, es obvio que existe un valor λ tal que:

(1) $$\nabla f + \lambda \nabla h = 0$$

Por ejemplo:



Suponemos que, en este caso, la curva de restricción es una circunferencia (anotamos como $g(x,y) = c$) y el valor de $f(x,y) = w$ varía entre 6 y 17. Geométricamente, estamos buscando un punto de la curva cuyo valor sea el máximo. Si empezamos con el valor 17, vemos que la curva no tiene ninguna intersección con la f, entonces de esta forma movemos hasta el valor 14 donde solo hay un punto de intersección (en caso general, podemos hacer movimiento hasta el momento en que la curva solo corta con las líneas de nivel en un punto), y llamamos a este punto, el punto P. Es evidente que este punto es el que consigue el valor máximo en la curva. En el punto P, ∇g es perpendicular a la recta tangente de la curva que pasa por el punto P. Y además ∇f es perpendicular a la recta tangente que pasa por cualquier punto de su línea de nivel, incluido obviamente el punto P. Como la curva y la línea de nivel son tangentes y nada más se corta en un solo punto, el punto P, sus rectas tangentes son paralelas, lo que indica que existe un valor real λ, tal que $\nabla f + \lambda \nabla g = 0$.

2.2 El lagrangiano

Vemos en el apartado anterior que ,para tratar un problema de optimización, debemos, al mismo tiempo, considerar una función objetiva y una serie de restricciones formadas por varias igualdades. De aquí nos surge la pregunta "¿se pueden combinar todas esas informaciones en una ecuación sola?". La respuesta es sí y, a esa ecuación que recopila todas informaciones, la llamamos el lagrangiano, que está definida como (aquí empezamos con un caso simple de dimensión dos):

$$L(x, y, \lambda) = f(x, y) + \lambda h(x, y)$$

Donde $f(x, y)$ es la función objetiva del problema y $h(x, y)$ la restricción de igualdad.

Notamos que

$$\nabla L = \begin{pmatrix} \frac{\partial f}{\partial x} & +\lambda \frac{\partial h}{\partial x} \\ \frac{\partial f}{\partial x} & +\lambda \frac{\partial h}{\partial x} \\ & h \end{pmatrix}^T = (\nabla f + \lambda \nabla h, h)$$

Eso nos dice el punto extremo, el valor máximo o minino de la función objetiva, $\nabla L = 0$, ya que se anularán al mismo tiempo los dos componentes.

Definición 2.1: El parámetro que aparece en esta ecuación, λ, se llama *el multiplicador de lagrange*. Y el método de aproximar al valor extremo de la función objetiva a través de la construcción del lagrangiano suponiendo que su gradiente es igual cero, se llama *el método del multiplicador de lagrange*.

Observación 2.2: Una vez transformamos el problema de optimización en la ecuación del lagrangiano correspondiente, nuestro objetivo ya no será encontrar el valor máximo o mínimo, sino encontrar el punto crítico (x, y, λ) del lagrangiano.

Ejemplo 2.3: Encontrar el valor extremo de la función $f(x, y) = xy$ sujeto a la restricción:

$$h(x, y) = x^2 + y^2 - 2 = 0$$

En primer lugar, construimos el lagrangiano y encontramos su gradiente:

$$L(x, y, \lambda) = xy + \lambda(x^2 + y^2 - 2)$$

$$\nabla L = \begin{pmatrix} y & +2\lambda x \\ \mathrm{x} & +2\lambda \mathrm{y} \\ x^2 + y^2 - 2 & \end{pmatrix}^T = 0$$

Lo que nos indica es que:

(2) $$y + 2\lambda x = 0$$

(3) $$x + 2\lambda y = 0$$

(4) $$x^2 + y^2 - 2 = 0$$

Eliminamos x con las ecuaciones (2) y (3):

$$y - 4\lambda^2 y = 0$$

Como $y \neq 0$, nos lleva a que:

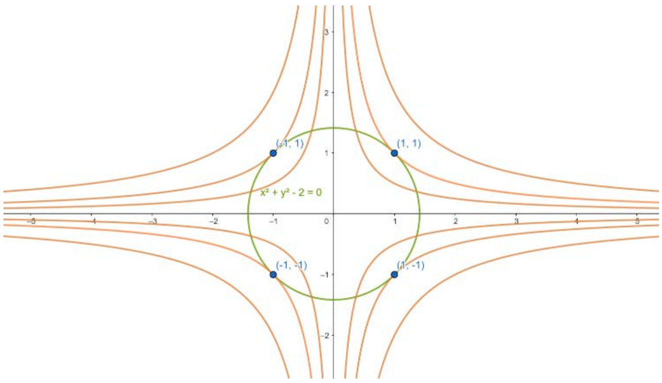
$$\lambda^2 = \frac{1}{4}$$

$$\lambda = \pm\frac{1}{2}$$

Como $x = \pm y$, por la ecuación (3), lo sustituimos en la (4), conseguimos:

$$x = \pm 1; y = \pm 1,$$

Entonces, hay 4 puntos extremos de la función f sujeta a la restricción h: (1, 1), (-1, -1), (-1, 1) y (1, -1). Sustituimos esos puntos de la función objetiva y podemos ver que los primeros dos puntos nos llevan al valor máximo y los dos últimos al valor mínimo de la función f.



2.3 Generalización en el caso de varias restricciones

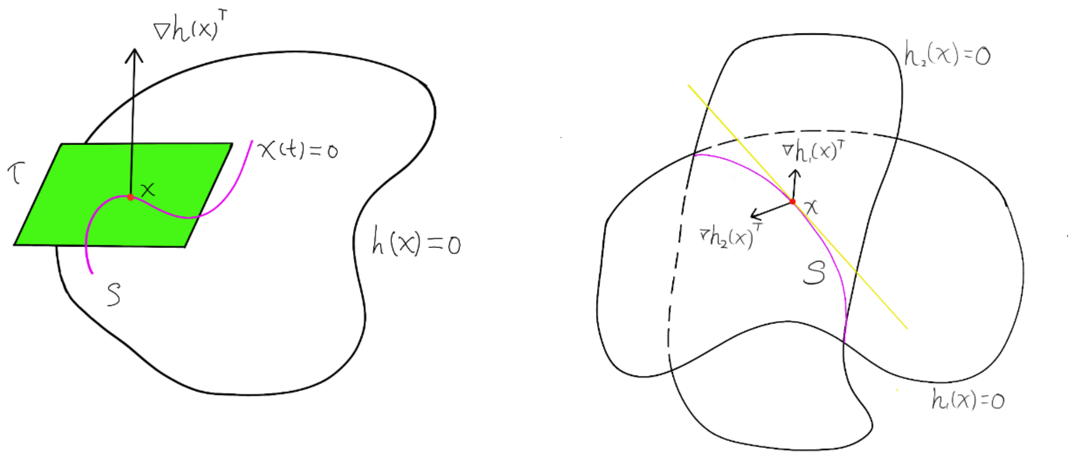
Ahora, nos enfocamos en el caso de un problema de optimización con múltiples restricciones de igualdades. Sea $h = (h_1, h_2, \cdots h_n)^T$ una función de restricción definida de $\mathbb{R}^n$ a $\mathbb{R}^m$. Consideremos el siguiente problema de optimización:

$$Minimiza\ f(x)$$

$$Sujeto\ a\ h(x) = 0$$

Cada función h_i define una superficie $s_i = \{x|h_i(x) = 0\}$ de dimensión n-1 en el espacio $\mathbb{R}^n$. Esta superficie es suave y diferenciable ya que cada h_i es diferenciable. Por lo tanto, si juntamos todas las restricciones, h_i , podemos tener una intersección de m superficies, y lo definimos como:

$\{x|\mathrm{h}(x) = 0\} = S = s_1 \cap s_2 \dots \cap s_m = \{x|h_1(x) = 0\} \cap \{x|h_n(x) = 0\} \dots \cap \{x|h_n(x) = 0\}$



(a) (b)

Figura 1: El espacio tangente en la superficie de la figura (a) viene definido por solo una restricción de igualdad h(x)=0 (plano de color verde) mientras que en la figura (b) existen dos restricciones, que son $h_1(x) = 0$ y $h_2(x) = 0$, cuya intersección es S y que, en este caso, el plano tangente sería la recta tangente de S, y aparece marcada en color amarillo en la figura.

La superficie S, en caso general, es una variedad diferencial de dimensión n-m. Consideramos un punto p en la superficie, definida de forma anterior, como intersección de todas restricciones, y todas las curvas que pasan por el punto p, y que están contenidas en el plano S. El espacio tangente τ consiste de todas las rectas tangentes de las curvas que pase por el punto p. Sea x(t) una de esas curvas, y x=p. Como x(t) está en todas las superficies S_i, cada vector tangente a S en el punto x es automáticamente un vector tangente de S_i. Eso implica que el espacio tangente τ a S en x es un subespacio de los τ_i a S_i en el mismo punto, para 1<=i<=m. Mientras, el gradiente ∇h_i es ortogonal a τ_i, porque:

$$h_i'(x) = \nabla h_i(x) \cdot x'(t) = 0$$

Entonces, ∇h_itiene que ser ortogonal al subespacio τ de los τ_i, $\nabla h_i \in \tau^{\perp}$.

Definición 2.4: Se dice que un punto P es regular, si los vectores gradientes de este punto son linealmente independientes.

Teorema 2.5: En un punto regular, x, de la superficie S definido con una serie de restricciones de igualdades $h(x) = 0$, el espacio tangente de este punto es:

$$\{y|\nabla h(x)y = 0\}$$

Donde:

$$\nabla h^T = (\nabla h_1 \cdots \nabla h_m)$$

La matriz de columna $\nabla h(x)$ está formada por los ∇h_i.

Lema 2.6: Deja $\bar{x}$ de ser un extremo local de la función f sujeto a las restricciones $h(x) = 0$. Entonces para todos y en el espacio tangente de la superficie formada por las restricciones en el punto $\bar{x}$, cumple:

$$\nabla f(\bar{x})y = 0$$

El lema nos indica que $\nabla f(\bar{x}) \in \tau^{\perp}$. Y de ahí, podemos decir que, en un punto extremo de la función, $\bar{x}$, ∇f es una combinación lineal de los ∇h_i, $1 \leq i \leq m$. Esto es válido solo para los puntos regulares. El siguiente teorema afirma que el método del multiplicador de Lagrange es una condición necesaria de la existencia de un punto extremo del problema de optimización.

Teorema 2.7 (condición necesaria del primer orden): Sea $\bar{x}$ un punto extremo local de la función f, sujeto a las restricciones $h(x) = 0$. Además, suponemos que $\bar{x}$ es un punto regular de la superficie formado por las restricciones, entonces existe un valor $\lambda \in \mathbb{R}^m$ tal que:

$$\nabla f(x) + \lambda^T \nabla h(\bar{x}) = 0$$

Dada la condición necesaria de primer orden junto con las restricciones, tenemos en total unos n+m ecuaciones de n+m incógnitas, los componentes de $\bar{x}$ y de λ. Por lo tanto, el sistema puede tener una única resolución, al menos localmente.

Ejemplo: Suponemos que tenemos una función objetiva de multivariable x, y, z, y queremos encontrar su valor máximo sujetando a una restricción, el problema está planteado de la forma siguiente:

$$\text{Maximizar } f(x, y, z) = xyz$$

$$\text{Sujeto a } h(x, y, z) = xy + yz + xz - \frac{C}{2} = 0$$

Donde $C > 0$ es un valor predefinido. Si consideramos el lagrangiano $L(x, y, z) = xyz + \lambda\left(xy + yz + xz - \frac{c}{2}\right)$, la condición necesaria del primer orden es fácil de encontrar:

(6) $$yz + \lambda(y + z) = 0$$

(7) $$xz + \lambda(x + z) = 0$$

(8) $$xy + \lambda(x + y) = 0$$

Y, junto con la restricción $h(x, y, z) = 0$, podemos sumar las tres ecuaciones alistadas arriba, y tenemos:

$$(xy + yz + xz) + 2\lambda(x + y + z) = 0$$

Si sustituimos la información de la restricción, tenemos:

$$\frac{C}{2} + 2\lambda(x + y + z) = 0$$

Por lo tanto, está claro que $\lambda \neq 0$, ni x, y, z pueden ser cero.

Para resolver las ecuaciones $(6) - (8)$, podemos multiplicar (6) por x, y (7) por y, y después podemos restar para obtener:

$$\lambda(x - y)z = 0$$

Hacemos lo mismo con otras ecuaciones y obtenemos:

$$\lambda(y - z)x = 0$$

Como ninguna de las variables puede tomar el valor 0,

$$x = y = z = \sqrt{\frac{C}{6}}$$

Es la única solución a la condición necesaria.

Además, podemos derivar la condición necesaria de segundo orden para los problemas de optimización asumiendo que f y h son dos veces diferenciables.

Teorema 2.8 (condición necesaria del segundo orden): Sea $\bar{x}$ un local mínimo de la función f, sujeto a las restricciones $h(x) = 0$. Además, suponemos que $\bar{x}$ es un punto regular de la superficie formada por las restricciones. Y denotamos con F, la matriz hessiana de f, y con H_i, la hessiana de h_i, para $1 \leq i \leq m$. Entonces existe un valor $\lambda \in \mathbb{R}^m$ tal que:

$$\nabla f(x) + \lambda^T \nabla h(\bar{x}) = 0$$

La matriz:

$$L(\bar{x}) = F(\bar{x}) + \sum_{i=1}^{n} \lambda_i \, H_i \, (\bar{x})$$

Es semidefinida positiva en el espacio tangente $\{y | \nabla h(\bar{x}) y = 0\}$.

Teorema 2.9 (condición suficiente del segundo orden): Sea $\bar{x}$ punto que satisfaga $h(\bar{x}) = 0$, y un valor λ tal que:

$$\nabla f(x) + \lambda^T \nabla h(\bar{x}) = 0$$

Sea también la matriz $L(\bar{x})$ definida anteriormente semidefinida positiva en el espacio tangente $\{y | \nabla h(\bar{x}) y = 0\}$. Entonces, $\bar{x}$ es estrictamente, el valor mínimo localmente de la función *f* sujeto a las restricciones $h(x) = 0$.

Ejemplo 2.10: Consideramos el problema

$$\text{Minimizar} \quad f(x) = x_1 x_2 + x_2 x_3 + x_1 x_3$$

$$\text{Sujeto a} \qquad x_1 + x_2 + x_3 = 3$$

La condición necesaria de primer orden se convierte en:

$$x_2 + x_3 + \lambda = 0$$

$$x_1 + x_3 + \lambda = 0$$

$$x_1 + x_2 + \lambda = 0$$

Podemos resolver estas tres ecuaciones junto con la restricción y obtenemos:

$$x_1 = x_2 = x_3 = 1 \quad \text{y} \quad \lambda = -2$$

Entonces $\bar{x} = (1,1,1)^T$

Ahora tenemos que observar la condición suficiente de segundo orden para determinar si el problema logra un valor máximo o mínimo en el punto $x_1 = x_2 = x_3 = 1$. Encontramos la matriz:

$$L(\bar{x}) = F(\bar{x}) + \lambda H(\bar{x}) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

En general esta matriz ni es definida positiva ni es definida negativa. Pero en el espacio tangente $M = \{y | y_1 + y_2 + y_3 = 0\}$, notamos que

$$y^T L y = y_1(y_2 + y_3) + y_2(y_1 + y_3) + y_3(y_1 + y_2)$$

$$= -(y_1^2 + y_2^2 + y_3^2) < 0 \qquad para\ todos\ y \neq 0$$

Entonces L es definida negativa en el espacio tangente M y la solución que hemos encontrado es un máximo, al menos localmente.

2.4 Restricción de desigualdad: Condiciones KKT (Karush-Kuhn-Tucker)

Para conseguir poca la intuición del problema de optimización con restricciones de desigualdades, podemos empezar con el caso más simple, que solo consta de una restricción de desigualdad. El problema puede ser plantado de la siguiente forma:

$$\text{Maximiza } f(x, y)$$

$$\text{Sujeto a } g(x, y) \leq 0$$

Como, en este problema, hemos sustituido la igualdad por una desigualdad en la restricción, este cambio nos surge un nuevo concepto fundamental.

Definición 2.11: Decimos que una restricción de desigualdad es activa en un punto $x \in \mathbb{R}^n$, si cumple que $g(x) = 0$. En caso contrario, si $g(x) < 0$, la llamamos inactiva. Por convención, en el caso de las restricciones de igualdades, siempre pensamos que son activas en los puntos factibles.

Teorema 2.12: Sea $\bar{x}$ una solución del problema planteado arriba: $\bar{x}$ maximiza f respecto a la restricción $g(\bar{x}) \leq 0$. Si $g(\bar{x}) = 0$, $\nabla g(\bar{x}) \neq 0$. Y en este caso, existe un multiplicador λ, tal que

$$\begin{gathered} \nabla L(\bar{x}, \lambda) = 0 \\ \lambda g(\bar{x}) = 0 \\ \lambda \geq 0 \\ g(\bar{x}) = 0 \end{gathered} \tag{9}$$

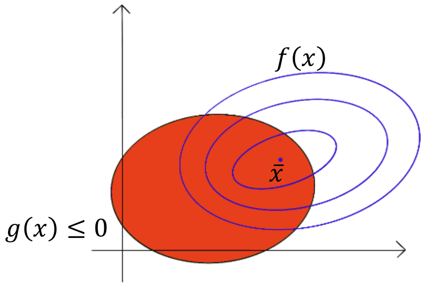
Donde $L(\bar{x}, \lambda) = f(\bar{x}) - \lambda g(\bar{x})$, es el lagrangiano del problema.

Observación 2.13: las condiciones que hemos alistado arriba, en el teorema, son conocidas como las condiciones KKT (Karush-Kuhn-Tucker).

Observación 2.14: cuando el problema consiste en minimizar la función objetiva, la condición (9) debe ser sustituida por $\lambda \leq 0$.

Casi una demostración 2.15: Aquí observamos la diferencia entre las restricciones activas e inactivas en el punto $\bar{x}$, discutimos dos casos.

Caso 1: la restricción es inactiva en el punto $\bar{x}$, es decir $g(\bar{x}) < 0$



Eso implica que el punto $\bar{x}$ es un valor máximo, pero está por adentro de la región limitada por la restricción. Entonces $\nabla f(\bar{x}) = 0$. En este caso, como la restricción no sirve exactamente para nada, la suponemos inactiva y le asignaremos el valor del multiplicador $\lambda = 0$. Así, las condiciones KKT se convierten:

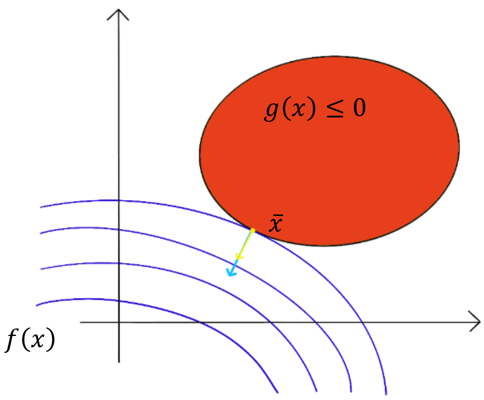
$$\nabla L(\bar{x}, \lambda) = \nabla f(\bar{x}) = 0$$

$$\lambda g(\bar{x}) = 0 \cdot g(\bar{x}) = 0$$

$$0 = \lambda \geq 0$$

$$g(\bar{x}) < 0$$

Caso 2: la restricción es activa en el punto $\bar{x}$, es decir $g(\bar{x}) = 0$. En este caso, la restricción pasará lo mismo como si fuera una restricción de igualdad.



Por lo tanto,

$$\nabla L(\bar{x}, \lambda) = \left(\frac{\partial L}{\partial \bar{x}}, \frac{\partial L}{\partial x}\right) = (0{,}0)$$

Entonces las condiciones KKT de nuevo, se convierten en:

$$\nabla L(\bar{x}, \lambda) = 0$$

$$\lambda g(\bar{x}) = 0 \cdot \lambda = 0$$

$$\lambda \geq 0$$

$$g(\bar{x}) = 0$$

Como $\bar{x}$ es el punto máximo, $\nabla f(\bar{x})$ y $\nabla g(\bar{x})$ debe estar en la misma dirección, lo cual nos implica que (19), $\lambda \geq 0$.

Ejemplo 2.16:

$$Maximiza \quad f(x,y) = xy$$

$$Sujeta\ a \quad g(x,y) = x^2 + y^2 - 2 \leq 0.$$

El lagrangiano de este problema es: $L(x,y,\lambda) = xy - \lambda(x^2 + y^2 - 2)$, y la condición KKT es:

$$\nabla L = (y - 2\lambda x, x - 2\lambda y) = (0,0)$$

$$\lambda(x^2 + y^2 - 2) = 0 \ donde\ \lambda \geq 0$$

$$g(x,y) = x^2 + y^2 - 2 \leq 0$$

Entonces, consideramos dos casos:

Caso 1, $\lambda = 0$, el sistema se convierte en:

$$(x,y) = (0,0)$$

$$x^2 + y^2 \leq 2$$

Entonces, obviamente, la única solución es $(0,0)$.

Caso 2. $g(x,y) = 2x + y - 2 = 0$, en este caso, la condición KKT es:

$$(y - 2\lambda x, x - 2\lambda y) = (0,0)$$

$$(x^2 + y^2 - 2) = 0$$

$$\lambda \geq 0$$

$$x^2 + y^2 \leq 2$$

Fácilmente, podemos obtener $\lambda = \frac{y}{2x} = \frac{x}{2y} \rightarrow x^2 = y^2$, juntamos esa información con la segunda ecuación y tendremos $\begin{cases} x^2 - y^2 = 0 \\ x^2 + y^2 = 2 \end{cases}$. La solución sería $x = \pm 1$ y $y = \pm 1$.

$$(x,y,\lambda) = \left(1,1,\frac{1}{2}\right)$$

$$(x,y,\lambda) = \left(-1,-1,\frac{1}{2}\right)$$

$$(x,y,\lambda) = \left(-1,1,-\frac{1}{2}\right)$$

$$(x,y,\lambda) = \left(1,-1,-\frac{1}{2}\right)$$

Descartamos los últimos dos, porque $\lambda \geq 0$, como $f(1,1) = f(-1,-1) = 1$, existen dos puntos máximos, que son (1,1) y (-1,-1).

Observación 2.17: Para el problema que minimiza la misma función objetiva, como se exige $\lambda \leq 0$, las dos soluciones que hemos descartado son valores mínimos.

2.5 Restricción de varias desigualdades

El planteamiento del problema es básicamente el mismo que en el apartado anterior:

$$\text{Maximiza } f(x)$$

$$\text{Sujeto a } g(x) \leq 0$$

La única diferencia es que $y = (g_1, g_2 \cdot \cdots g_k)$ y $x \in \mathbb{R}^n$. Exactamente lo mismo que antes, cuando una restricción $g_i(x) = 0$, la llamamos activa y, en caso contrario, inactiva.

Teorema 2.18: Sea $\bar{x} \in \mathbb{R}^n$, es un punto que maximiza nuestro problema, y también suponemos que las primeras k_0 restricciones de desigualdades son activas, mientras las restas $k - k_0$ son inactivas. Además, suponemos que la matriz jacobiana formada por las restricciones activas es de rango k_0 o, en otras palabras, el vector gradientes de esas restricciones activas es linealmente independiente.

Consideramos el lagrangiano:

$$L(x, \lambda) = f(x) - \lambda^T g(x)$$

Donde $\lambda \in \mathbb{R}^k$, en este caso, podemos construir la condición KKT de la siguiente forma:

$$\nabla L(x, \lambda) = 0$$

$$\lambda^T g(x) = 0$$

$$\lambda \geq O_k$$

$$g(x) \leq O_k$$

Observación 2.19: El teorema enunciado anteriormente puede ser reformulado de la siguiente forma:

Sea $\bar{x} \in \mathbb{R}^n$, una solución que maximza la función objetiva $f(x)$ sujeto a $g(x) \leq 0$, sean las primeras k_0 componentes de g restricciones activas, es decir,

$$\left(g_1(\bar{x}), g_2(\tilde{x}) \cdot \cdots g_{k_0}(\bar{x})\right) = O_{k_0}$$

Además, suponemos que $\nabla g_1(\bar{x}), \ldots, \nabla g_{k_0}(\bar{x})$ son linealmente independientes. Entonces:

$$\nabla f(\bar{x}) \in < \nabla g_1(\bar{x}), \ldots, \nabla g_{k_0}(\bar{x}) >$$

Podemos encontrar un $\lambda \in \mathbb{R}^k$ tal que sus primeros k_0 componentes cumplan:

$$\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_{k_0} \nabla g_{k_0}(\bar{x})$$

$$\lambda_1 \geq 0, \cdots \lambda_{k_0} \geq 0$$

Los demás $k - k_0$ componentes son todos nulos.

Ejemplo 2.20:

$$\text{Maximiza } f(x_1, x_2) = -(x_1 - 4)^2 - (x_2 - 4)^2$$
$$\text{sujeto a } g_1(x_1, x_2) = x_1 + x_2 - 4 \leq 0,$$
$$g_2(x_1, x_2) = x_1 + 3x_2 - 9 \leq 0.$$

El lagrangiano puede ser escrito como:

$$L(x_1, x_2) = -(x, -4)^2 - (x_2 - 4)^2 - \lambda_1(x_1 + x_2 - 4) - \lambda_2(x_1 + 3x_2 - 9).$$

La condición KKT sería:

$$-2(x_1 - 4) - \lambda_1 - \lambda_2 = 0$$
$$-2(x_2 - 4) - \lambda_1 - 3\lambda_2 = 0$$
$$\lambda_1(x_1 + x_2 - 4) = 0$$
$$\lambda_2(x_1 + 3x_2 - 9) = 0$$
$$\lambda_1 \geq 0, \lambda_2 \geq 0$$
$$x_1 + x_2 - 4 \leq 0, x_1 + 3x_2 - 9 \leq 0$$

Si analizamos la tercera y la cuarta condición $\lambda_1(x_1 + x_2 - 4) = 0$, $\lambda_2(x_1 + 3x_2 - 9) = 0$, podemos obtener los siguientes casos:

Caso 1: $g_1(x_1, x_2) = g_2(x_1, x_2) = 0$. Las dos restricciones son activas y podemos tener

$$x_1 = \frac{3}{2}, x_2 = \frac{5}{2}, \lambda_1 = 6, \lambda_2 = -1$$

Pero como $\lambda_2 \geq 0$, entonces este caso no nos da ninguna resolución correspondiente a la KKT.

Caso 2: g_1 es activa mientras g_2 no lo es, que implica $\lambda_2 = 0$. En este caso podemos tener

$$x_1 = x_2 = 2, \lambda_1 = 4, \lambda_2 = 0$$

Y este sí que es la solución de KKT.

Caso 3. g_2 es activa mientras g_1 no lo es, que implica $\lambda_1 = 0$. En este caso podemos tener

$$x_1 = 3.3, x_2 = 1.8$$

Pero esta solución viola la restricción $x_1 + x_2 - 4 \leq 0$, entonces tampoco nos da ninguna solución correspondiente a la KKT.

Caso 4: Las dos restricciones son inactivas, que implica $\lambda_1 = \lambda_2 = 0$. Podemos tener $x_1 = x_2 = 4$, lo cual viola la restricción $x_1 + x_2 - 4 \leq 0$, por lo que no hay ninguna solución en este caso.

La solución que nos ha dado para el caso 2, es una solución de KKT, pero no podemos decir que es el valor que maximiza la función, porque KKT solo es una condición necesaria. Aún se debe verificar a través de la matriz hessiana. Para más detalles, se puede consultar el apartado anterior "la condición suficiente de segundo orden".

2.6 Restricciones mixta

El último apartado de este tema se trata del caso más complejo del problema de optimización, que incluye tanto las igualdades como las desigualdades y las restricciones. El problema puede ser planteado de la siguiente forma:

$$\text{Maximiza } f(x) \text{ donde } x \in \mathbb{R}^n$$

$$\text{Sujeto a } g(x) \leq O_k$$

$$h(x) = O_m$$

Teorema 2.21: Sea $\bar{x} \in \mathbb{R}^n$, es un punto que maximiza nuestro problema, y también suponemos que las primeras k_0 restricciones de desigualdades son activas, mientras las restas $k - k_0$ son inactivas. Además, suponemos que la matriz jacobiana formada por las restricciones activas y las m restricciones de igualdades es de rango $k_0 + m$. En otras palabras, el vector gradientes de esas restricciones activas junto con las restricciones de igualdades son linealmente independientes en el punto $\bar{x}$.

Consideramos el lagrangiano:

$$L(x, \lambda) = f(x) - \lambda^T g(x) - \mu^T h(x)$$

Donde $\lambda \in \mathbb{R}^k$, $\mu \in \mathbb{R}^m$, en este caso, podemos construir la condición KKT de la siguiente forma:

$$\nabla L(\bar{x}, \lambda) = 0$$

$$\lambda^T g(\bar{x}) = 0$$

$$h(\bar{x}) = O_m$$

$$\lambda \geq O_k$$

$$g(x) \leq O_k$$

Observación 2.22: El teorema enunciado anteriormente puede ser reformulado de la siguiente forma:

Sea $\bar{x} \in \mathbb{R}^n$, una solución que maximiza la función objetiva $f(x)$ sujeto a $g(x) \leq 0$ y $h(x) = 0$, sean las primeras k_0 componentes de g restricciones activas, es decir,

$$\left(g_1(\bar{x}), g_2(\tilde{x}) \cdot \cdots g_{k_0}(\bar{x})\right) = O_{k_0}$$

Además, suponemos que $\nabla g_1(\bar{x}), \dots, \nabla g_{k_0}(\bar{x})$, $\nabla h_1(\bar{x}), \dots \nabla h_m(\bar{x})$ son linealmente independientes. Entonces:

$$\nabla f(\bar{x}) \in < \nabla g_1(\bar{x}), \dots, \nabla g_{k_0}(\bar{x}), \nabla h_1(\bar{x}), \dots \nabla h_m(\bar{x}) >$$

Así, podemos encontrar un $\mu \in \mathbb{R}^m$ y un $\lambda \in \mathbb{R}^k$ tal que sus primeros k_0 componentes cumplen:

$$\nabla f(\bar{x}) = \lambda_1 \nabla g_1(\bar{x}) + \cdots + \lambda_{k_0} \nabla g_{k_0}(\bar{x}) + \mu_1 \nabla h_1(\bar{x}) + \cdots + \mu_m \nabla h_m(\bar{x})$$

$$\lambda_1 \geq 0, \cdots \lambda_{k_0} \geq 0$$

Los demás $k - k_0$ componentes son todos nulos.

Capítulo 3. Teoría de portafolio

La Teoría de la Cartera Moderna (MPT, por sus siglas en inglés), o análisis de varianza media, es un marco matemático para armar una cartera de activos de tal manera que el rendimiento esperado se maximice para un nivel dado de riesgo.

Es una formalización y extensión de la diversificación en la inversión, la idea de que poseer diferentes tipos de activos financieros es menos riesgoso que poseer solo un tipo. Su idea clave es que el riesgo y el rendimiento de un activo no deben evaluarse por sí mismos, sino por la forma en que contribuyen al riesgo y rendimiento general de una cartera. Utiliza la variación de los precios de los activos como un *proxy* del riesgo. El economista Harry Markowitz presentó el MPT en un ensayo de 1952, por el cual más tarde recibió el Premio Nobel de Economía.

En este apartado, se pretende resolver el problema de selección de un portafolio óptimo como un problema de optimización, aplicando el método de multiplicador de Lagrange.

3.1. Conceptos básicos

Un instrumento de inversión que se puede comprar y vender con libertades normalmente se llama *activo*. Suponemos que algún día compramos un activo por x_0 euros en una fecha concreta y, después de un periodo, lo vendemos por x_1 euros. Podemos llamar la proporción definida

$$R = \frac{x_1}{x_0}$$

como el retorno de nuestra inversión. Naturalmente, la tasa de retorno de esta viene definida como

$$r = \frac{x_1 - x_0}{x_0} = R - 1$$

Entonces,

$$x_1 = Rx_0 \qquad y \qquad x_1 = (1 + r)x_0$$

Todo el mundo sabe que cuando poseemos un activo, lo podemos vender cuando queremos pero, ¿ y si no lo tenemos aún? ¿Se puede vender algo que aún no es tuyo? ¡La respuesta es sí! Podemos vender algún activo que todavía no es nuestro gracias al mundo financiero. Llamamos este procedimiento como la *venta corta* y funciona así. Supongamos que tú quieres corta (venta corta) una acción particular XXX. Te diriges al *bróker* que te asiste para saber si él tiene la misma acción en la piscina de activos que mantiene para sus clientes. Si él posee algo de la acción XXX, tú le prestas una cierta cantidad asegurando que se la vas a devolver la misma cantidad algún día del futuro. Desde entonces, en tu lista de activos se marca como cantidad negativa de la acción XXX, lo cual indica que tú debes a tu bróker tantas cantidades de acciones. En la mayoría de los casos, no se apunta el valor total de la deuda sino solo la cantidad de acciones que debes. Como en el tiempo inicial ya has vendido tantas cantidades de acciones XXX que has prestado de tu bróker por x_0 euros, cuando toca la fecha de vencimiento debes comprar la misma cantidad de acciones XXX para devolvérselo, sea en este momento, el coste es x_1. Obviamente, si $x_1 < x_0$, es decir que el precio de acción XXX ha bajado a lo largo del tiempo, por lo tanto, por esta operación te han generado un ingreso de $x_0 - x_1$ euros y, en caso contrario, te causa una pérdida. (Como el valor de una acción solo puede bajarse hasta 0 euros como mínimo, pero sí que puede subirse hasta un valor extramental, lo que implica, la pérdida de la venta corta no está acotada y puede ser enorme.)

En una operación de venta corta, el retorno y la tasa de retorno se define de la siguiente forma:

$$R = \frac{-x_1}{-x_0} = \frac{x_1}{x_0} \quad y \quad r = \frac{(-x_1) - (-x_0)}{-x_0} = \frac{x_1 - x_0}{x_0}$$

Ahora, supongamos que tenemos un portafolio formado por n activos de riesgo y tenemos un presupuesto inicial de x_0 euros para invertir en este portafolio. La cantidad de dinero que vamos a asignar al activo i está definida por $x_{0i} = wx_0$, donde w_i es el peso que ocupa el activo i en el portafolio. Aquí, el peso puede ser negativo en el caso de venta corta. De todas formas, debemos siempre asegurar que la suma del peso de todos los activos es igual a 1, es decir, $\sum_{i=1}^{n} w_i = 1$, y además, para asegurar, no sobrecargar nuestro presupuesto, debemos suponer que:

$$\sum_{i=1}^{n} w_i x_0 = x_0 \sum_{i=1}^{n} w_i = x_0$$

Si con R_i denotamos el retorno del activo i, entonces transcurrido un periodo, podemos conseguir:

$$x_1 = \sum_{i=1}^{n} R_i w_i x_0 = x_0 \sum_{i=1}^{n} R_i w_i$$

Y, por lo tanto, el retorno global del portafolio sería:

$$R = \sum_{i=1}^{n} R_i w_i$$

Además, tenemos que la ratio de retorno del activo i es $r_i = R_i - 1$, donde $i = 1,2, \dots, n$. La ratio global en este caso, será:

$$r = R - 1 = \left(\sum_{i=1}^{n} R_i w_i \right) - \sum_{i=1}^{n} w_i = \sum_{i=1}^{n} (R_i - 1) w_i = \sum_{i=1}^{n} r_i w_i$$

3.2. Teoría de portafolio Markowitz

El apartado anterior nos ha enseñado los conceptos básicos de la teoría de portafolio, de donde vemos que, normalmente, existe un número infinito de formas para asignar el peso a cada activo contenido en un portafolio. Pero entre ellos, ¿cuál es la mejor opción? Para encontrar su resolución, convendría aclarar un poco más sobre el concepto "mejor": puede ser el portafolio con el mayor retorno, con el menor riesgo o un equilibrio entre estos dos aspectos.

La ratio de retorno de un portafolio puede ser modelizado como una variable aleatoria y su riesgo, o la volatilidad, obviamente, será la varianza de este variable aleatoria, ya que la varianza representa la dispersión de los datos.

Sea r_i la variable aleatoria asociada a la ratio de retorno del activo i, para $i = 1,2,\dots,n$, definimos el vector aleatorio:

$$z^T = (r_1, r_2, \dots, r_n)$$

Sea $\mu_i = E(r_i)$, $\mathrm{cov}(z) = \Sigma$, donde $\Sigma \in M_{nxn}$ definida de la forma siguiente:

$$\Sigma = \left(\sigma_{ij}\right)_{1 \le i,j \le n}$$

donde $\sigma_{ii} = var(r_i)$ y $\sigma_{ij} = cov\left(r_i, r_j\right)$. Si denotamos el peso asignado a cada activo como un vector $w = \left(w_1, w_{2,\dots,} w_n\right)^T$, entonces la ratio de retorno de este portafolio es igual a $r = \sum_{i=1}^{n} r_i w_i$, y r también es una variable aleatoria con esperanza igual a $E(r) = m^T w$, y varianza $var(r) = w^T \Sigma w$.

3.3. La selección de portafolio como un problema de optimización

Como se ha mencionado en el apartado anterior, la selección de portafolio puede convertirse en un problema de optimización de tres formas distintas.

a. Minimiza el riesgo (la varianza o la volatilidad)

Sea μ_b el menor retorno que podemos aceptar para esta inversión, como un individuo racional, siempre pretendemos asumir el mínimo riesgo posible, entonces podemos modelizar esta situación como el problema de optimización de la forma siguiente:

$$minimiza: \qquad \frac{1}{2}var(r) = \frac{1}{2}w^T\Sigma w$$

$$sujeto\ a: \ E(r) = m^T w - \mu_b \geq 0 \ donde\ e^T w = 1$$

El vector e tiene todos sus componentes igual a 1.

Vemos que el lagrangiano de este problema sería:

$$L(w,\lambda,\gamma) = \frac{1}{2}w^T\Sigma w - \lambda(m^T w - \mu_b) - \gamma(e^T w - 1)$$

donde $\lambda, \gamma \in \mathbb{R}$

Según la idea transmitida con el tema anterior, podemos formar las condiciones KKT como las siguientes:

(1) $$\frac{\partial L}{\partial w} = 0 = \Sigma w - \lambda m - \gamma e$$

(2) $$\mu_b \leq m^T w, e^T w = 1, 0 \leq \lambda$$

(3) $$\lambda(m^T w - \mu_b) = 0$$

Como en este programa cuadrático tenemos la matriz Σ simétrica, ya que $\sigma_{ij} = \sigma_{ji}$, y positivamente definida, si (w,λ,γ) es un triple que satisfaga la condición KKT, entonces w es necesariamente una solución del problema de optimización. De hecho, podemos ver fácilmente que este problema de minimización es factible, entonces siempre existe una solución y que cumple las condiciones KKT.

El problema puede ser clasificado en dos casos respecto a la (3) de la condición KKT.

- $\lambda = 0$: En este caso la restricción de desigualdad es inactiva[2], es decir, esta restricción podría ser eliminada ya que no nos aporta ninguna información útil. En este caso, las condiciones KKT se reducen a solo dos ecuaciones:

[2] Consultar el contenido de la sección 2.4 “Restricciones de desigualdades”.

$$\Sigma \mathrm{w} - \gamma e = 0$$
$$e^T \mathrm{w} = 1$$

Si multiplicamos los dos lados de la primera ecuación por Σ^{-1}, tendremos $\mathrm{w} = \gamma \Sigma^{-1} e$, después sustituimos este resultado a la segunda ecuación, $e^T \gamma \Sigma^{-1} e = 1$. De ahí podemos conseguir el valor del parámetro $\gamma = (e^T \Sigma^{-1} e)^{-1}$ y, por lo tanto:

$$\mathrm{w} = (e^T \Sigma^{-1} e)^{-1} \Sigma^{-1} e$$

Hasta ahora, ya tenemos el triple que cumple las condiciones KKT, $(w, \lambda, \gamma) = ((e^T \Sigma^{-1} e)^{-1} \Sigma^{-1} e, 0, (e^T \Sigma^{-1} e)^{-1})$. Entonces w es una solución del problema de optimización.

Observación 3.1: Como en este caso la restricción de desigualdad es inactiva, el valor mínimo que hemos conseguido también se reconoce como el peso de varianza mínima global, ya que no depende de la restricción, solo de la función objetiva y lo denotamos por $w_{\min -var}$.

Una vez encontrado este valor $w_{\min -var}$, debemos verificar si cumple la restricción $m^T w \geq \mu_b$. Si lo cumple, ya no tenemos que hacer nada más, ya que este peso es lo asociado a la varianza mínima de forma global. En caso contrario, debemos suponer el segundo caso, que la restricción de desigualdad sí que es activa.

- $\lambda \neq 0$ y esto implica que $m^T w = \mu_b$: multiplicamos la (1) de las condiciones KKT por Σ^{-1}, tendremos:

  (4) $$w = \lambda \Sigma^{-1} m + \gamma \Sigma^{-1} e$$

  Si sustituimos este resultado a la condición (2), podemos tener dos ecuaciones:

  $$\mu_b = \lambda m^T \Sigma^{-1} m + \gamma m^T \Sigma^{-1} e$$
  $$1 = \lambda m^T \Sigma^{-1} e + \gamma e^T \Sigma^{-1} e$$

  O equivalentemente, lo podemos reescribir en la forma más compacta:

  (5) $$\begin{bmatrix} m^T \Sigma^{-1} m & m^T \Sigma^{-1} e \\ m^T \Sigma^{-1} e & e^T \Sigma^{-1} e \end{bmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = \begin{pmatrix} \mu_b \\ 1 \end{pmatrix}$$

  La matriz:

  $$T = \begin{bmatrix} m^T \Sigma^{-1} m & m^T \Sigma^{-1} e \\ m^T \Sigma^{-1} e & e^T \Sigma^{-1} e \end{bmatrix} = \begin{pmatrix} m \\ e \end{pmatrix} \Sigma^{-1} (m\ e)$$

  es siempre semidefinida positiva. Para verificar esta propiedad, solo tenemos que ver si su determinante es estrictamente positivo, es decir:

  $$0 < |T| = (m^T \Sigma^{-1} m)(e^T \Sigma^{-1} e) - (m^T \Sigma^{-1} e)^2$$

  Esta propiedad siempre se cumple si m y e son vectores linealmente independientes.

  - Si m y e son linealmente dependiente, es decir que existe $\tau \in \mathbb{R}$ tal que $m = \tau e$, y que nos implica que $|T| = 0$. En este caso, si $\mu_b \neq \tau$, el problema es necesariamente no factible. Mientras que si $\mu_b = \tau$, $w_{\min -var}$ resuelve el problema, podemos calcular primero su valor y después verificar si este valor cumple la restricción.

- Si m y e son linealmente independientes, nos implica que $|T| > 0$, en este caso, aplicamos el método de Cramer al sistema de (5) y tendrá la solución:

$$\lambda = \frac{\mu_b(e^T\Sigma^{-1}e) - e^T\Sigma^{-1}m}{|T|} \quad y \quad \gamma = \frac{m^T\Sigma^{-1}m - \mu_b(m^T\Sigma^{-1}e)}{|T|}$$

Entonces, lo simplificamos:

$$\lambda = e^T v \quad y \quad \gamma = -m^T v$$

Donde

$$v = |T|^{-1}\Sigma^{-1}(\mu_b e - m)$$

Sustituimos el valor λ y γ a (4) y tendremos:

$$\begin{aligned} w &= e^T v\Sigma^{-1}m - m^T v\Sigma^{-1}e \\ &= \frac{\mu_b(e^T\Sigma^{-1}e)-e^T\Sigma^{-1}m}{|T|}(\Sigma^{-1}m)\text{-}\frac{m^T\Sigma^{-1}m-\mu_b(m^T\Sigma^{-1}e)}{|T|}(\Sigma^{-1}e) \\ &= \frac{\mu_b(e^T\Sigma^{-1}e)(e^T\Sigma^{-1}m)-(e^T\Sigma^{-1}m)^2}{|T|}\frac{\Sigma^{-1}m}{e^T\Sigma^{-1}m} + \\ &\quad \frac{-(m^T\Sigma^{-1}m)(e^T\Sigma^{-1}e)+\mu_b(m^T\Sigma^{-1}e)(e^T\Sigma^{-1}e)}{|T|}\frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e} \\ &= \alpha\frac{\Sigma^{-1}m}{e^T\Sigma^{-1}m} - (\alpha-1)\frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e} \\ &= \alpha w_{mk} - (\alpha-1)w_{min-var} \\ &= (1-\alpha)w_{min-var} + \alpha w_{mk} \end{aligned}$$

Donde

$$w_{mk} = \frac{\Sigma^{-1}m}{e^T\Sigma^{-1}m}$$

Lo denotamos como el peso del mercado y

$$\alpha = \frac{\mu_b(m^T\Sigma^{-1}e)(e^T\Sigma^{-1}e) - (m^T\Sigma^{-1}e)^2}{|T|}$$

Observación 3.2: el peso óptimo que resuelve este problema de optimización puede ser representado como una combinación lineal de dos pesos, $w_{min-var}$ y w_{mk}. El peso del mercado incorpora todas las informaciones de los activos que son operativos.

Algoritmo 3.3:

El procedimiento del cálculo puede ser resumido en los siguientes pasos:

<u>Verificación de la factibilidad</u>

En este paso, solo debemos verificar si m y e son vectores linealmente independientes. Si son linealmente dependientes, existe un $\tau \in \mathbb{R}$ tal que $m = \tau e$.

En este caso, el problema no es factible si $\mu_b > \tau$. Si $\mu_b \leq \tau$, entonces calculamos Σ^{-1} y evaluamos el peso $w_{min-var}$. Este peso puede resolver el problema.

<u>Verificación de la solución</u>

Si el problema es factible, entonces calculamos Σ^{-1}, y evaluamos el peso $w_{min-var}$ de la siguiente forma:

$$w_{min-var} = \frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e}$$

Si $m^T w_{\min-var} \geq \mu_b$, $w_{\min-var}$ es la solución del problema.

<u>Cálculo de la solución de dos portafolios</u>

Si el problema es factible, pero $w_{\min-var}$ no es su solución, entonces debemos calcular el peso del mercado de la siguiente forma:

$$w_{mk} = \frac{\Sigma^{-1}m}{e^T\Sigma^{-1}m}$$

Y formamos el vector

$$v = w_{mk} - w_{min-var}$$

La solución del problema puede ser expresada como:

$$w = w_{\min-var} + \alpha v$$

Para determinar el valor de α, usamos la igualdad $m^T w = \mu_b$ para conseguir

$$\alpha = \frac{\mu_b - m^T w_{min-var}}{m^T v}$$

b. Maximiza el retorno (la rentabilidad)

El problema puede ser modelizado de la siguiente forma:

$$Maximiza\colon E(r) = m^T w$$

$$sujeto\ a: \frac{1}{2}var(r) = \frac{1}{2}w^T\Sigma w - \frac{1}{2}\sigma_b \leq 0\ donde\ e^T w = 1$$

El lagrandiano puede ser escrito como:

$$L(w,\lambda,\gamma) = m^T w - \lambda\left(\frac{1}{2}w^T\Sigma w - \frac{1}{2}\sigma_b\right) - \gamma(e^T w - 1)$$

Análogamente, podemos formalizar las condiciones KKT:

$$\frac{\partial L}{\partial w} = 0 = m - \lambda\Sigma w - \gamma e$$

$$w^T\Sigma w - \sigma_b \leq 0,\ e^T w = 1, 0 \leq \lambda$$

$$\lambda\left(\frac{1}{2}w^T\Sigma w-\frac{1}{2}\sigma_b\right)=0$$

- Si la restricción es inactiva, es decir, $\lambda=0$, las condiciones KKT se convierten en:
$$m=\gamma e$$
$$e^Tw=1$$
La primera ecuación nos dice que m es paralelo a e, que significa que todos los activos del portafolio tienen la misma rentabilidad. Entonces, sea cual sea nuestra selección, la ratio de retorno se fijará en γ. Para poder minimizar el riesgo, debemos invertir todos nuestros recursos en el activo que tiene menos riesgo, es decir, $w=(0{,}0,\ldots,1{,}0,\ldots,0)$

- Si la restricción es activa, implica que $\lambda\neq 0$, $w^T\Sigma w=\sigma_b$. En este caso, las condiciones KKT serían:

(1) $$m=\lambda\Sigma w+\gamma e$$
(2) $$w^T\Sigma w=\sigma_b$$
(3) $$e^Tw=1$$
(4) $$\lambda>0$$

De la condición (1), podemos conseguir
$$w=\frac{1}{\lambda}\Sigma^{-1}m-\frac{\gamma}{\lambda}\Sigma^{-1}e$$
Y lo sustituimos a la (2) y (3):

(4) $$\frac{1}{\lambda}e^T\Sigma^{-1}m-\frac{\gamma}{\lambda}e^T\Sigma^{-1}e=1=\frac{b}{\lambda}-\frac{a\gamma}{\lambda}$$
(5) $$\frac{1}{\lambda^2}(m^T\Sigma^{-1}m)-\frac{2\gamma}{\lambda}(e^T\Sigma^{-1}m)+\frac{\gamma^2}{\lambda^2}(e^T\Sigma^{-1}e)=\sigma_b=\frac{1}{\lambda^2}\mathrm{a}-\frac{2\gamma}{\lambda}\mathrm{b}+\frac{\gamma^2}{\lambda^2}c$$

De (4), despejamos las variables, podemos tener: $\gamma=\frac{b-\lambda}{a}$, y lo sustituimos a (5):

(6) $$\left(a+\frac{b^2c}{a^2}\right)\frac{1}{\lambda^2}-\left(\frac{2b^2}{a}+\frac{2bc}{a^2}\right)\frac{1}{\lambda}+\left(\frac{2b}{a}+\frac{c}{a^2}\right)=\sigma_b$$

(6) es una ecuación de segundo grado de $\frac{1}{\lambda}$, lo resolvemos, y sea λ_1,λ_2 las dos soluciones de la ecuación (en caso $\Delta<0$, no existe solución real, entonces el problema no es factible), destacamos los valores negativos, ya que $\lambda>0$ estrictamente. Sustituimos la solución, en caso de existir, a la (4) para evaluar γ. Y por último, los sustituimos a la (1) para conseguir el w, que sería el peso de la cartera que nos da la rentabilidad máxima bajo con el riesgo menor que σ_b.

c. Un equilibrio de los dos aspectos (Teorema de dos fondos)

En muchos casos, nuestra intención no es conseguir una rentabilidad máxima ni un riesgo mínimo, sino encontrar un equilibrio entre estos eso dos aspectos. O sea, queremos encontrar una estrategia, con la cual siempre podamos invertir con el menor riesgo posible para conseguir una rentabilidad concreta y con la mayor rentabilidad asumiendo una cierta cantidad de riesgo.

Para resolver este problema de optimización, podemos modelizar el caso de la siguiente forma:

$$minimiza\text{:}\ \frac{1}{2}var(r) - E(r) = \frac{1}{2}w^T\Sigma w - \lambda m^T w$$

$$sujeto\ a:\ e^T w = 1$$

En la función objetiva, con el parámetro $\lambda \geq 0$, el término $-\lambda m^T w$ disminuye su valor de la función con el argumento de la rentabilidad mientras que, cuanto menor es el término $w^T\Sigma w$, el valor de la función se vuelve menor también. Tenemos el lagrangiano del problema:

$$L(w,\lambda,\gamma) = \frac{1}{2}w^T\Sigma w - \lambda m^T w - \gamma(e^T w - 1)$$

Y las condiciones KKT asociadas son:

(1) $$\frac{\partial L}{\partial w} = \Sigma w - \lambda m - \gamma e = 0$$

(2) $$e^T w = 1$$

Vemos que la condición (1) es igual que la condición (1) del problema a) solo que ahora ya no nos exige un valor mínimo μ_b de la rentabilidad. Sustituimos la condición (2) a la (1) y conseguimos:

$$\gamma = \frac{1 - \lambda m^T \Sigma^{-1} e}{e^T \Sigma^{-1} e}$$

Y w_λ, el peso del portafolio en función del parámetro λ es igual a:

$$w_\lambda = \frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e} + \alpha\left[\frac{\Sigma^{-1}m}{m^T\Sigma^{-1}e} - \frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e}\right]$$

$$= (1-\alpha)\frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e} + \alpha\frac{\Sigma^{-1}m}{m^T\Sigma^{-1}e}$$

$$= (1-\alpha)w_{min-var} + \alpha w_{m_k}$$

Donde $w_{min-var}$ y w_{m_k} ya ha sido definido anteriormente, y

$$\alpha = \alpha_\lambda = \lambda(m^T\Sigma^{-1}e)$$

En este caso la rentabilidad μ_λ sería:

$$\mu_\lambda = m^T w_\lambda = (1-\alpha)m^T w_{min-var} + \alpha m^T w_{m_k}$$

$$= m^T w_{min-var} + \alpha m^T (w_{m_k} - w_{min-var})$$

$$= \mu_{min-var} + \lambda (m^T \Sigma^{-1} e)(w_{m_k} - w_{min-var})$$

$$= \mu_{min-var} + \lambda \frac{\delta}{e^T \Sigma^{-1} e}$$

Donde $\delta = (m^T \Sigma^{-1} m)(e^T \Sigma^{-1} e) - (m^T \Sigma^{-1} e)^2 = |T|$

Entonces, cuando λ toma el valor 0, $\mu_0 = \mu_{min-var}$, este es el caso de que no fijemos ninguna rentabilidad mínima y nos satisface con la rentabilidad del menor riesgo del mercado, es decir, solo pretender que nuestro riesgo sea lo mínimo posible (Mínimo absoluto del mercado).

Observación 3.4: Vemos cuando λ se va creciendo, tendiendo hacia al infinito, que la rentabilidad asociada, μ_λ, también tenderá a ∞ si $\delta = |T| > 0$, es decir, cuando m y e son linealmente independientes. Entonces, si el λ va corriendo desde 0 hasta infinito, su rentabilidad asociada, μ_λ, se va corriendo desde $\mu_0 = \mu_{min-var}$ hasta infinito. Por lo tanto, para visualizar mejor la situación, solo tenemos que trazar el gráfico de la curva:

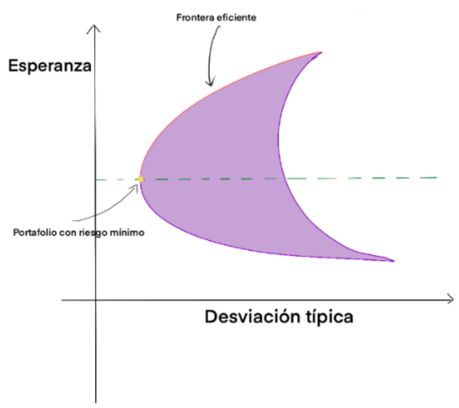
$$\left(\sqrt{(w_\lambda^T \Sigma w_\lambda)}, r_\lambda\right) = \left(\sqrt{var(r_\lambda)}, E(r_\lambda)\right)$$

Donde:

$$r_\lambda = w_\lambda^T r$$

Eso es, la construcción de un sistema de coordenadas, poniendo la desviación típica (el riesgo) en el eje X, y la esperanza (la rentabilidad) en el eje Y.

Definición 3.5: Según el teorema de portafolio de Markowitz, esta curva se llama la curva de eficiencia o la frontera eficiente. Un portafolio, asociado a un punto de la curva, se llama portafolio eficiente.



Para aclarar más sobre el concepto, imaginamos una región formada por los puntos, (Riesgo, rentabilidad), asociados a todos los portafolios que podamos formar. Eso es, una región de los puntos de forma $\left(\sigma_i = \sqrt{var(r_i)}, \mu_i = E(r_i)\right)$, para $i = 1,2,\ldots,n$, asociada a todos w_i, portafolios, que cumple $e^T w_i = 1$ y la frontera superior es justo la curva de eficiencia que estamos buscando.

Teorema de dos fondos 3.6: Bajo ciertas condiciones, las condiciones KKT, un portafolio óptimo de cualquier inversor, (depende de su preferencia de riesgo y de rentabilidad), puede construirse manteniendo cada uno de los dos portafolios en proporción adecuada.

Demostración 3.7: Dados dos valores μ_1 *y* μ_2, que satisface $\mu_{\min -var} < \mu_1 < \mu_2$, podemos encontrar el portafolio w_i, correspondiente a cada uno de ellos, resolviendo las condiciones KKT, formado de la siguiente forma:

(1) $$0 = \Sigma w_i - \lambda_i m - \gamma_i e$$

(2) $$\mu_i = m^T w_i$$

(3) $$e^T w = 1$$

Es fácil de resolver este sistema de ecuaciones lineales. Sea $(w_i, \lambda_i, \gamma_i)$, para $i = 1,2$, las soluciones que estamos buscando. Entonces, podemos definir:

$$(w_\alpha, \lambda_\alpha, \gamma_\alpha) = (1-\alpha)(w_1, \lambda_1, \gamma_1) + \alpha(w_2, \lambda_2, \gamma_2)$$

Obviamente, $\mu_\alpha = (1-\alpha)\mu, + \alpha\mu_2$ para cualquier valor $\alpha \in \mathbb{R}$, entonces el portafolio asociado se va definido como $r_\alpha = (1-\alpha)r_1 + \alpha r_2$. Como α se va corriendo de $\frac{\mu_{min-var} - \mu_1}{\mu_2 - \mu_1}$ hasta $+\infty$, el punto asociado, $(\sqrt{Var(r_\alpha)}, E(r_\alpha))$ se va corriendo a través de la curva de eficiencia entre los portafolios, r_1y r_2.

Observación 3.8: Hemos visto en el caso a) de 2.4, que todos los portafolios eficientes pueden ser expresados como una combinación lineal de dos portafolios concretos, $w_{min-var}$ y w_{m_k}, sea $w_1 = (1-\alpha)w_{min-var} + \alpha w_{m_k}$ y $w_2 = (1-\beta)w_{min-var} + \beta w_{m_k}$. Dado un portafolio eficiente cualquier, μ_c, este siempre puede ser expresado como:

$$w_c = (1-\theta)w_{\min -var} + \theta w_{mk}$$
$$= (1-\theta, \theta)\begin{bmatrix} 1-\alpha & \beta \\ 1-\beta & \beta \end{bmatrix}^{-1} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

3.4. Teorema de diversificación de Markowitz

Hemos visto en el apartado anterior que cualquier portafolio eficiente puede ser expresado como una combinación lineal de dos portafolios eficientes, pero, ¿cómo elegimos estos portafolios para que su combinación tenga el menor riesgo posible?

Supongamos que tenemos dos portafolios w_1 y w_2 , además $a, b \in \mathbb{R}$, dos parámetros reales y el portafolio definido como, $w = aw_1 + bw_2$, tiene su esperanza $\mu = a\mu_1 + b\mu_2$ y su varianza ve definida como

(1) $\quad Var(\mathrm{r}) = Var(ar_1 + br_2) = a^2Var(r_1) + b^2Var(r_2) + 2abCov(r_1, r_2)$

como la correlación $\rho_{r_1,r_2} = \frac{Cov(r_1,r_2)}{\sigma_1\sigma_2}$. Entonces, si podemos encontrar dos portafolios tales que su correlación sea igual a -1, (1) se convierte en:

$$Var(\mathrm{r}) = (a\sigma_1 - b\sigma_2)^2$$

Donde $\sigma_1^2 = var(r_1)\ y\ \sigma_2^2 = var(r_2)$

En este caso, si encontramos a, b tal que $a\sigma_1 = b\sigma_2$, el portafolio w no tiene nada de riesgo.

Sea w_1, w_2 dos portafolios tal que la correlación entre ellos es igual a -1, para conseguir un portafolio con riesgo nulo, solo debemos encontrar los parámetros a, b tal que:

$$a\sigma_1 = b\sigma_2$$

$$a + b = 1$$

Resolvemos el sistema para tener: $a = \frac{\sigma_2}{\sigma_1+\sigma_2}$, $b = \frac{\sigma_1}{\sigma_1+\sigma_2}$ y el portafolio formado tiene su retorno $\mu = \frac{\sigma_1\mu_1+\sigma_2\mu_2}{\sigma_1+\sigma_2}$.

Este fenómeno de correlación negativa se llama el fenómeno de Markowitz, y es una de las ideas básicas de la diversificación para los inversores: para encontrar un portafolio con menor riesgo, se deben encontrar dos portafolios eficientes con correlación lo más cercana posible a -1.

Esa idea puede extenderse a una combinación de varios portafolios (o activos). Sea $r_1, r_2, \ldots, r_n$, n portafolios **no correlacionados**, con $Var(r_i) \leq C\ para\ i = 1,2, \ldots, n$, donde C es un constante. Entonces:

$$Var(\bar{r}) = Var(d^T r) = \sum_{i=1}^{n} d_i^2 Var(r_i) \leq C \sum_{i=1}^{n} d_i^2$$

sea $d_i = \frac{1}{n}$, vemos que:

$$\lim_{n\to\infty} var(\bar{r}) = \frac{c}{\infty} \to 0$$

Observación 3.9: El caso anterior, nos indica que para un portafolio formado por una serie de portafolios (o activos), cuanto más grande es n, asumirá el menor riesgo.

Como normalmente, no podemos esperar a encontrar uno e incluso solo dos portafolios (o activos) que tengan correlación -1. En el caso general, sea $\bar{r} = d^T r_i$ un portafolio formado por n portafolios r_i, entonces su varianza es igual a:

$$Var(\bar{r}) = Var(d^T r) = \sum_{i=1}^{n} d_i^2 Var(r_i) + \sum_{i,j=1,i\neq j}^{n} d_i d_j Cov(r_i, r_j)$$

Supongamos que $d_i = \frac{1}{n}$, entonces,

$$\sum_{i=1}^{n} d_i^2 Var(r_i) = \left(\frac{1}{n}\right)^2 n \frac{1}{n} \sum_{i=1}^{n} Var(r_i) = \frac{1}{n}\widetilde{Var}$$

Donde $\widetilde{Var}$ representa el promedio de las varianzas de los portafolios (o activos).

Además:

$$\sum_{i,j=1,i\neq j}^{n} d_i d_j Cov(r_i, r_j) = \left(\frac{1}{n}\right)^2 (n^2 - n)\widetilde{Cov}$$

Donde $\widetilde{Cov} = \frac{1}{n^2-n} \sum_{i,j=1,i\neq j}^{n} Cov(r_i, r_j)$ es el promedio de la covariancia.

$$Var(\bar{r}) = \frac{1}{n}\widetilde{Var} + (1 - \frac{1}{n})\, \widetilde{Cov}$$

Evidentemente, si $\widetilde{Var} \leq C$

$$\lim_{n\to\infty} Var(\bar{r}) = \widetilde{Cov}$$

De este resultado podemos concluir que, para un n suficiente grande, el riesgo se iguala al promedio de la covariancia. En caso de poder encontrar los productos que no sean correlacionados, el portafolio tiene su riesgo nulo. Pero, en la vida real, esto es casi imposible, ya que normalmente, todos productos del mercado tienen correlación positiva entre ellos, y además tampoco podemos esperar a incluir casi infinitos productos en nuestra cartera.

Definición 3.10: $\widetilde{Cov}$, el promedio de la covarianza de todos los productos de la cartera se llama riesgo sistemático y es el riesgo que no podemos diversificar de ninguna forma y también el menor riesgo que podemos alcanzar si invertimos en productos en la proporción adecuada.

3.5 La inclusión de un activo sin riesgo

En los apartados anteriores, nuestros portafolios estaban formados por activos con riesgo si ahora incluimos en el portafolio un activo sin riesgo puede ser un bono, una obligación, un fondo de inversión monetario, o cualquier activo que tiene su riesgo casi nulo. Sea r_f es el variable aleatorio asociado a este activo. Actualmente, tenemos un portafolio de forma, $r = \left(r_f, x^T\right)^T$, donde, $x = (r_1, r_2 \cdots r_n)^T$, los n activos con riesgo. Como suponemos que r_f tiene riesgo nulo, la covarianza entre este activo y los demás igual a 0. Tenemos la matriz de covariación:

$$\Sigma' = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma \end{bmatrix}$$

Donde Σ es la matriz de covariación del vector x. Ahora, el programa cuadrático de Markowitz puede ser escrito como:

$$minimiza\text{:}\ \frac{1}{2} var(r) = \frac{1}{2} w^T \Sigma w$$

$$sujeto\ a:\ E(r) = m^T w + r_f w_0 \geq \mu_b\ donde\ e^T w + w_0 = 1$$

Aquí w_0 denota el peso del activo sin riesgo y μ_b, igual que en los casos anteriores, denota la rentabilidad mínima aceptable. Como nuestra rentabilidad es siempre mayor que r_f, ya que, en caso contrario, solo invertimos en el activo sin riesgo, suponemos $\mu_b \geq r_f$. Además, suponemos que Σ, la matriz de covarianza de x, no es singular.

El lagrangiano en este caso es:

$$L(w_0, w, \lambda, \gamma) = \frac{1}{2} w^T \Sigma w + \lambda(m^T w + r_f w_0 - \mu_b) + \gamma(e^T w + w_0 - 1)$$

Las condiciones KKT en este caso puede ser escrito como:

(1) $$0 = \lambda r_f + \gamma$$

(2) $$0 = \Sigma w - \lambda m - \gamma e$$

(3) $$0 = \lambda\left(r_f w_0 + m^T w - \mu_b\right)$$

(4) $$1 = w_0 + e^T w$$

(5) $$\mu_b \leq r_f w_0 + m^T w$$

(6) $$\lambda \geq 0$$

Igual que en los casos anteriores, debemos discutir si la condición (5) es activa o no.

- Si (5) es inactivo, $\lambda = 0$, y la condición (1) nos dice que, en este caso, $\gamma = 0$ también. Por lo siguiente, la condición (2) implica que $w = 0$. Como $w_0 + w = 1$, $w_0 = 1$. Entonces, podemos concluir que cuando la condición (5) es inactiva, el portafolio óptimo consiste solo en el activo sin riesgo.

- Si (5) es activo, lo cual nos implica $\mu_b = r_f w_0 + m^T w$. Podemos aplicar la misma técnica que hemos usado en 2.3 caso a). En primer lugar, multiplica (2) por Σ^{-1} y sustituimos γ por (1), tenemos:
  (7) $$w = \lambda \Sigma^{-1}(m - r_f e)$$
  Transformamos (4) y (5) para tener:
  $$1 - w_0 = e^T w = \lambda e^T \Sigma^{-1}(m - r_f e)$$
  $$\mu_b - r_f w_0 = m^T w = \lambda m^T \Sigma^{-1}(m - r_f e)$$
  Y lo podemos escribir como un producto matricial:
  $$\begin{bmatrix} 1 & e^T \Sigma^{-1}(m - r_f e) \\ r_f & m^T \Sigma^{-1}(m - r_f e) \end{bmatrix} \begin{pmatrix} w_0 \\ \lambda \end{pmatrix} = \begin{pmatrix} 1 \\ r_f \end{pmatrix}$$
  Resolvemos el sistema con la eliminación Gaussiana,
  $$\begin{pmatrix} w_0 \\ \lambda \end{pmatrix} = \begin{pmatrix} 1 - (\mu_b - r_f)\dfrac{e^T \Sigma^{-1}(m - r_f e)}{(m - r_f e)^T \Sigma^{-1}(m - r_f e)} \\ \dfrac{(\mu_b - r_f)}{(m - r_f e)^T \Sigma^{-1}(m - r_f e)} \end{pmatrix}$$
  - Si $m - r_f e = 0$, $w_0 = 1$ y $w = 0$
  - Si $m - r_f e \neq 0$, sustituimos el λ por la expresión (7), podemos tener:
    $$\begin{pmatrix} w_0 \\ w \end{pmatrix} = (1 - \alpha)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha \left[\begin{pmatrix} 1 - e^T \Sigma^{-1}(m - r_f e) \\ \Sigma^{-1}(m - r_f e) \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]$$
    Donde
    $$\alpha = \frac{(\mu_b - r_f)}{(m - r_f e)^T \Sigma^{-1}(m - r_f e)}$$
    Entonces, en este caso, el portafolio optimo es una combinación lineal de
    $$w_f = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ y } w_M = \begin{pmatrix} 1 - e^T \Sigma^{-1}(m - r_f e) \\ \Sigma^{-1}(m - r_f e) \end{pmatrix}$$

Esta solución nos lleva al siguiente teorema:

Teorema de un fondo 3.11: Si nuestro portafolio está construido por un activo sin riesgo y un fondo de inversión, ciertas cantidades de activos con riesgo, entonces, cualquier portafolio eficiente puede ser expresado como una combinación lineal del activo sin riesgo y este fondo de inversión.

Conclusión

Tras la redacción y estudio de este trabajo, soy capaz de aplicar una de las herramientas más potentes, los multiplicadores de Lagrange y su extensión (cuando se incluyen las desigualdades en las restricciones), las condiciones KKT, para resolver los problemas de optimización.

Por otro lado, tomando como una de sus aplicaciones la selección de las carteras de activos, no solo permite lograr una mejor compresión sobre esas herramientas, sino que también ofrece la posibilidad de crear una cartera óptima dependiendo de los requisitos. Se desprende del estudio, entonces, que cualquier cartera eficiente de activos con riesgo puede expresarse como una combinación de dos carteras eficientes dadas en proporción correcta y, en caso de la inclusión de un activo sin riesgo a la cartera, la necesidad de dos carteras eficientes se reduce a una sola.

La búsqueda de información y de las aplicaciones adecuadas, sobre todo para trazar los gráficos, ha sido vital, consultando artículos, trabajos, blogs y libros de muchos autores diferentes, para ver diferentes puntos de vista de estos resultados. Uno de los principales problemas con que me he encontrado ha sido, a la hora de desarrollar unos cálculos, ya que muchos de ellos son originales. Pero al final, los supero con el apoyo de mi tutor.

El desarrollo del trabajo me ha ayudado a profundizar mi conocimiento de Matemáticas y de Economía cuantitativa. Además, este trabajo me permite comprender mejor la vinculación entre ambas.

Bibliografía

[1] Luenberger, D. G. (1984). Linear and nonlinear programming Addison-Wesley. Reading, MA.

[2] Gavin, H. P., & Scruggs, J. T. (2012). Constrained optimization using lagrange multipliers. CEE 201L. Duke University. Constrained Optimization. Joshua Wilde, edition, August 13, 2013

[3] Trench, W. F. (2013). The Method of Lagrange Multipliers. S. Traductor, Trad.)

[4] Jia, Y. B. (2008). Lagrange Multipliers.

[5] Bramanti, M., Pagani, C. D., & Salsa, S. (2004). Calcolo infinitesimale e algebra lineare. Seconda edizione. Zanichelli.

[6] Burke, J. V. (2017). Markowitz Mean-Variance Portfolio Theory. Nonlinear Optimization.

[7] Markowitz, H. (1952). Portfolio selection. The journal of finance, 7(1), 77-91.

[8] Persson, J., Lejon, C., & Kierkegaard, K. (2007). Practical Application of Modern Portfolio Theory.

[9] Marling, H., & Emanuelsson, S. (2012). The Markowitz Portfolio Theory. November, 25, 2012.

[10] Barrera, P. S. (2008). Optimización de Portafolios.

[11] Shiryaev, A. N. (1999). Essentials of stochastic finance. World Scientific: Singapore.