Teoría del Portafolio Empírica: Dificultades en su Implementación

Carlos Rodríguez
PUCP

Diciembre 2018

Resumen

1. Enfoque Media-Varinza original de Markowitz
2. Dificultades en su implementación (Estimadores Muestrales)
3. Solución a los problemas de los estimadores muestrales
4. El modelo de Black-Litterman
5. Aplicación en Python

La raíz del problema

En el problema original de Markowitz, derivamos el vector de pesos óptimos del problema del portafolio , el cual es una función de:

$$w^* = f\left(c, \mu, \Sigma\right)$$

Como μ *y* ∑ no son directamente observables, tienen que ser estimados.

El enfoque más sencillo es usar el estimador muestral (promedio histórico) y reemplazarlos en la solución del problema de Markowitz,

Si "enchufamos" (plug-in) los estimadores muestrales de la media y la matriz de covarianzas μ y Σ en la solución de Markowitz, obtendremos un vector estimado

de los verdaderos pesos óptimos:

$$w^* = f\left(c, \mu, \Sigma\right)$$

La raíz del problema

El error de estimación de los parámetros μ y Σ se transfiere a los pesos óptimos del portafolio. Por lo tanto, los pesos óptimos serán diferentes a los pesos óptimos verdaderos.

El estimador muestra de la media y las covarinzas está dado por:

$$\widehat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_{t+1}$$

$$\widehat{\Sigma} = \frac{1}{T-N-2}\sum_{t=1}^{T}(r_{t+1}-\hat{\mu})(r_{t+1}-\hat{\mu})'$$

Sin embargo, en la práctica estos estimadores son rara vez utilizados, ya que se ha demostrado que los pesos del portafolio calculados usando los momentos muestrales son bastante imprecisos para tamaños de muestra realistas.

Consecuencias de usar los estimadores muestrales

Estos problemas se exacerban cuando el número de activos incluídos en el portafolio es grande.

Asimismo, a medida que aumenta el número de activos del portafolio, el número de elementos únicos en la matriz de Var-Cov aumenta a una tasa cuadrática.

Si tenemos N activos, entonces la matriz de Covarianzas tiene N(N+1)/2 términos únicos. Es decir, si tenemos 500 activos, se tendrían 500x(501)/2 = 125,250 elementos únicos en la matriz de covarianzas.

Como los pesos óptimos del portafolio w* dependen de todos estos parámetros, los pesos óptimos estimados tendrán mucho error.

Se ha mostrado que los pesos óptimos del portafolio son bastante sensibles a los inputs (retornos esperados y matriz de covarianzas). Es decir, pequeños cambios en estos inputs pueden producir cambios grandes en los pesos del portafolio, llevandonos a resultados poco realistas e inestables.

Método Delta para cálculo del error de estimación

Pero supongamos que usamos los momentos muestrales:
Qué puede estar pasando? La media muestral no es el estimador insesgado?

Pero el hecho que sean insesgado no quiere decir que sea el mejor estimador. Cómo sabemos que un estimador es el mejor estimador? Con insesgadez y por eficiencia (no sirve tener un estimador insesgado con mucha varianza)

El problema es el siguiente: estimar las medias (de los retornos) de largo plazo es muy difícil, pues tiene una gran varianza (el estimador de la media de los retornos no es eficiente)

Usando el método Delta y la teoría asintótica, podemos derivar la varianza del vector de pesos óptimos :

$$\widehat{w^*} \sim N\left(w(\widehat{\Sigma},\widehat{\mu}), \begin{bmatrix} w_\Sigma(\widehat{\Sigma},\widehat{\mu}) & w_\mu(\widehat{\Sigma},\widehat{\mu}) \end{bmatrix} \begin{bmatrix} \mathrm{var}(\widehat{\Sigma}) & 0 \\ 0 & \mathrm{var}(\widehat{\mu}) \end{bmatrix} \begin{bmatrix} w_\Sigma(\widehat{\Sigma},\widehat{\mu}) \\ w_\mu(\widehat{\Sigma},\widehat{\mu}) \end{bmatrix}\right)$$

Los estimados de la varianza de w son bastante inestables: tienen mucha varianza (no son eficientes)

Experimento de Jobson & Korkie

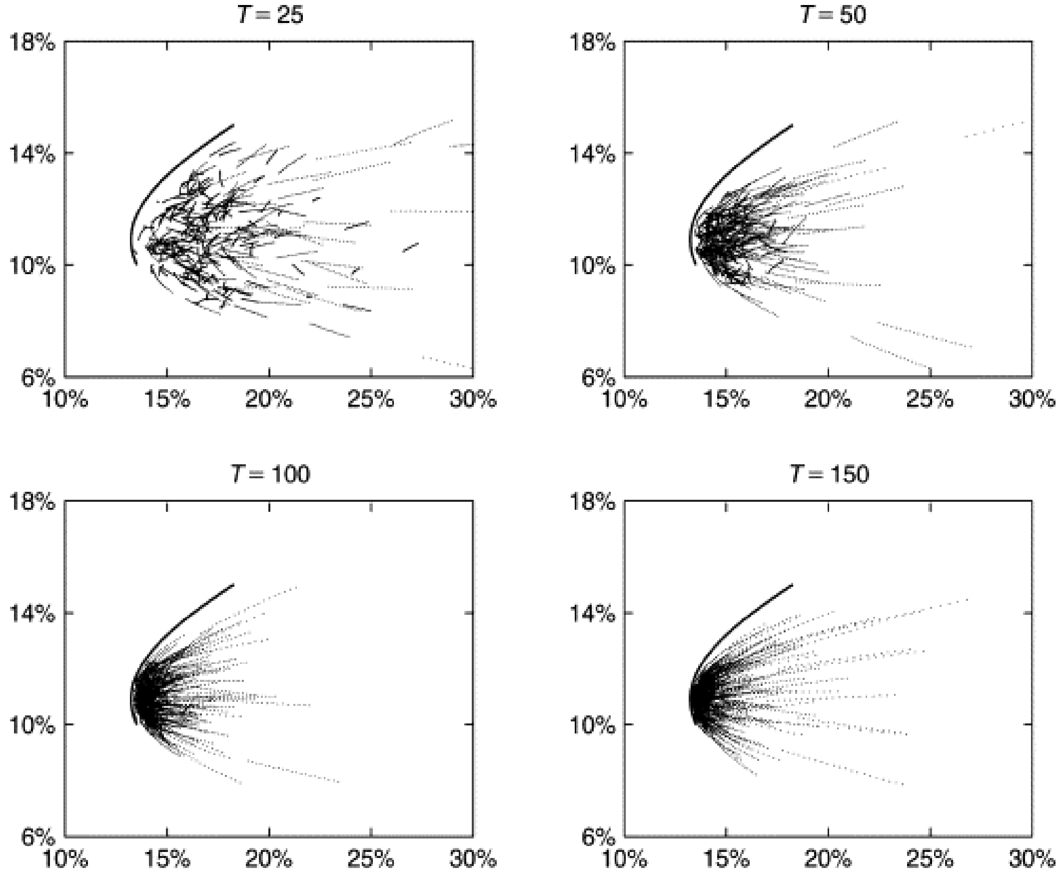
Realizaron un famoso experimento en el que asumieron que conocían cuál era la verdadera distribución de los Retornos, usando simulaciones de Montecarlo.

En la realidad, no conocemos la distribución completa (o el proceso generador de datos) sino que solo tenemos datos históricos de los retornos.

Jobson y Korkie construyeron la frontera eficiente de los agentes que solo han visto una parte de la historia y no han visto la distribución completa (usando diferentes tamaños de muestra)

Resultados Experimento de Jobson & Korkie

Para T=25 (25 años de datos) , la frontera eficiente estimada de los agentes cae muy adentro de la verdadera frontera eficiente



Incluso con los retornos históricos de 150 años, no te acercas a la verdadera frontera eficiente.

Las soluciones para el problema

Las soluciones pasan por obtener mejores estimadores (i.e más eficientes) que los estimadores muestrales usados convencionalmente.

Algunos métodos candidatos para obtener una mejor estimación estadística de los retornos esperados (primas por riesgo), incluyen por ejemplo:

i) Estimación por contracción (shrinkage estimation)
ii) Incorporación de modelos económicos y paradigma bayesiano: enfoque de **Black-Litterman**

Por el lado de la matriz de covarianzas, una posible solución es darle una estructura particular a la matriz. Esto da lugar a una parametrización parsimoniosa de la matríz de covarianzas. Para darle una estructura a dicha matriz, se usa frecuentemente:
-Modelos con Factores
-Index Models

Estos factores pueden ser obervables o extraídos de los retornos usando alguna técnica estadística como PCA.

Estimación por contracción (shrinkage estimation)

Es la solución más sencilla de implementar. Este método halla nuevos estimadores para μ y $\sum$.

Habíamos visto que los momentos muestrales eran estimadores insesgados pero no eficientes, por lo tanto, ahora vamos a buscar ahora estimadores de μ y $\sum$ que sean más eficientes, para lo cual este método usa el estimador de James y Stein (1961):

$$\mu_s = \delta\mu_0 + (1-\delta)\overline{\mu}$$

- Es decir, el método de Contracción estimama los retornos esperados como un promedio ponderado de la media muestral $\overline{\mu}$ y una constante μ_0 (valor benchmark)

- μ_0 puede ser el consenso de la rentabilidad que te dicen los analistas del mercado

- Se dice que "contraemos" la media muestral de los retornos $\overline{\mu}$ hacia una constante μ_0. El objetivo de este método es reducir el error de estimación, y eso lo hace "poniéndole un contrapeso" a la propia historia (datos muestrales)

Estimación por contracción (shrinkage estimation)

La matriz de varianzas y covarianzas también se puede estimar usando la misma idea:

$$\Sigma_s = \delta\widehat{S} + (1-\delta)\widehat{\Sigma}.$$

Es decir, como un promedio ponderado del estimador muestral y de un valor benchmark.

Lo interesante de este método es cómo hallar δ:

James y Stein plantearon que para reducir la varianza se podía permitir un poco de sesgo, plantearon entonces un problema de optimización de la combinación óptima de sesgo y varianza minimizando el error cuadrático medio: $(\mu - \mu_0)' \, \Sigma^{-1} \, (\mu - \mu_0)$

Obteniendo:

$$\delta^* = \min\left[1, \frac{(N-2)/T}{(\overline{\mu} - \mu_0)' \, \Sigma^{-1} \, (\overline{\mu} - \mu_0)}\right].$$

Enfoque Bayesiano

Ahora vamos a ver dos enfoques son diferentes:

1) Estimación bayesiana de los parámetros μ y $\sum$:

Se estima μ y $\sum$ con métodos bayesianos, y esos estimadores se sustituyen en la solución de pesos óptimos de portafolio, obteniendo la estimación bayesiana del portafolio de Markowitz.

2) Portafolios Bayesianos

Aca vamos a resolver un problema de optimización que toma en cuenta que el individuo no conoce la verdadera distribución de los parámetros (los retornos y las covarianzas). En Markowitz asumíamos que la distribución de los retornos dependen únicamente de μ y $\sum$ y que además los conocemos. Pero en la vida real, no conocemos ni la distribución ni el valor verdadero de estos parámetros. Entonces, tomando en cuenta esta falta de información, se reformula el problema de optimización. (para tomar en cuenta la incertidumbre de los parámetros)

Enfoque Bayesiano

En el enfoque bayesiano queremos obtener la distribución posterior o predictiva de los parámetros, la cual es proporcional a la multiplicación de la verosimilitud de los datos y los prios:

$$p(\theta|y) \propto p(y|\theta)p(\theta).$$

Vamos a tener algo interesante: que para ciertos casos (cuando usamos priors informativos), las formulas para los portafolios bayesianos van a ser una suerte de promedio ponderado. (tal como era en el método de contracción) entre el estimador muestral y nuestras distribuciones a priori.

El enfoque bayesiano aborda el riesgo de estimación desde una perspectiva conceptual diferente. En lugar de considerar a los verdaderos parámetros como fijos, los considera aleatorios. Las creencias del inversionista (conocimiento a priori) acerca de los parámetros de entrada (inputs), en combinación con los datos observados, producen una distribución completa de los retornos predichos que toma en cuenta explícitamente la incertidumbre de estimación y predictiva.

El enfoque de selección de portafolio de Black-Litterman

- Este modelo se ha convertido en la aplicación más prominente de la metodología de selección de portafolio bayesiana.
- Este enfoque es bastante usado en la práctica porque:

1. **El modelo permite incorporar las creencias (views) que tienen los Portafolio Managers sobre los retornos esperados sobre todos los activos de su portafolio o sobre un subconjunto de ellos (K<N).**

En el enfoque clásico de media-varianza se requería estimar las medias y covarianzas de todos los posibles activos del portafolio (osea monitorear $\mu\, y \sum$ para miles de activos financieros), lo cual resulta una tarea poco práctica.

Asimismo, los Portafolio Managgers usualmente solo tienen conocimiento y experiencia para proprocionar pronósticos sobre los Retornos de unos cuantos activos.

Por esta razón, no suelen usar en la práctica el enfoque media-varianza tradicional.

El enfoque de selección de portafolio de Black-Litterman

2. En el enfoque tradicional de Markowitz se obtienen soluciones de «esquina», en donde solo a unos pocos activos se les asigna pesos distintos de cero (es decir, se obtienen soluciones poco realistas y esto contradice los postulados originales de Markowitz, pues no se diversifica)

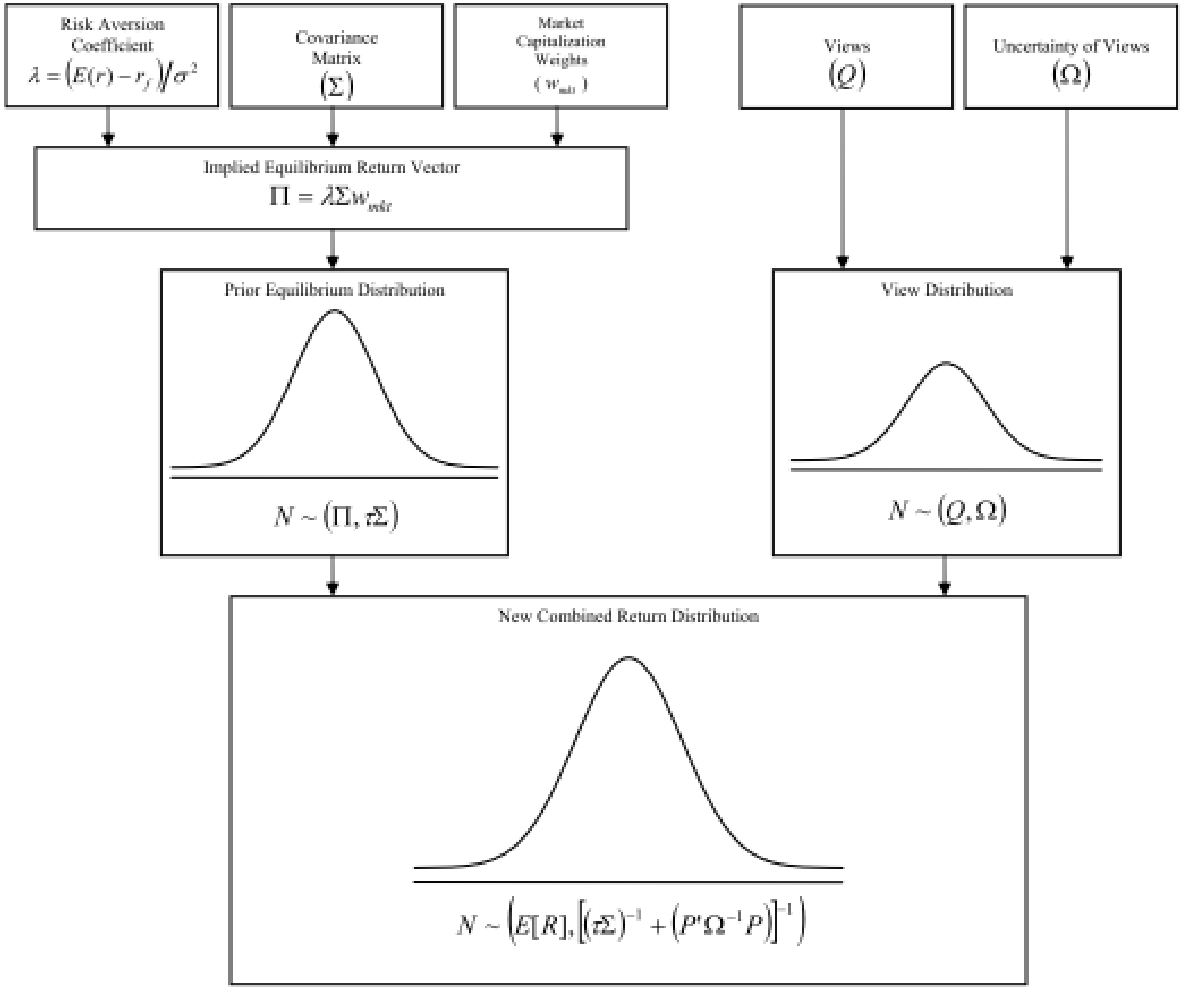
3. El enfoque bayesiano de la sección de Portafolio, y en particular el modelo de Black-Litterman, toman en cuenta la incertidumbre en la estimación.

En el modelo de BL, si no se expresan las «creencias» del inversionista sobre los Rendimientos Esperados de algunos activos, estos se centran en los «Rendimientos Esperados de Equilibrio»

Los métodos bayesianos son comúnmente críticados por la «arbitrariedad» involucrado en la elección de los priors. Pero, el modelo de BL ayuda a superar esta crítica ya que utiliza un modelo de valorización de activos de equilibrio (el modelo CAPM) como punto de referencia.

El Modelo de Black-Litterman

Figure 1 Deriving the New Combined Return Vector ($E[R]$)



* *The variance of the New Combined Return Distribution is derived in Satchell and Scowcroft (2000).*

El Modelo de Black-Litterman

El input mas importante en el enfoque de optimización de media-varianza es el vector de Retornos Esperados.

En búsqueda de un punto inicial razonable para los Retornos Esperados, Black y Literman (1992) usan los retornos de "equilibrio" como un punto neutral de partida.

Los **Retornos de equilibrio** son el conjunto de retornos que limpian el mercado. Los retornos de equilibrio se derivan usando un método de optimización inversa, en donde el vector de excesos de retorno de equilibrio implícitos se extraen de información conocida, usando la siguiente formula:

$$\Pi = \lambda \Sigma w_{mkt}$$

Donde:

- Π is the Implied Excess Equilibrium Return Vector (N x 1 column vector)
- λ is the risk aversion coefficient;
- Σ is the covariance matrix of excess returns (N x N matrix); and,
- w_{mkt} is the market capitalization weight (N x 1 column vector) of the assets.[4]

Ejemplo de Cálculo de los Retornos de Equilibrio

$$\Pi = \lambda \Sigma w_{mkt}$$

$$\begin{bmatrix} 0.08\% \\ 0.95\% \\ 3.95\% \\ 5.37\% \\ 5.14\% \\ 3.68\% \\ 6.12\% \\ 3.50\% \end{bmatrix} = 3.37 \begin{bmatrix} .0014 & .0015 & -.0008 & -.0017 & -.0010 & -.0007 & -.0015 & -.0006 \\ .0015 & .0076 & .0026 & -.0006 & -.0013 & -.0003 & -.0002 & .0005 \\ -.0008 & .0026 & .0251 & .0292 & .0208 & .0147 & .0248 & .0134 \\ -.0017 & -.0006 & .0292 & .0663 & .0359 & .0244 & .0490 & .0268 \\ -.0010 & -.0013 & .0208 & .0359 & .0468 & .0283 & .0520 & .0260 \\ -.0007 & -.0003 & .0147 & .0244 & .0283 & .0252 & .0314 & .0215 \\ -.0015 & -.0002 & .0248 & .0490 & .0520 & .0314 & .0809 & .0411 \\ -.0006 & .0005 & .0134 & .0268 & .0260 & .0215 & .0411 & .0276 \end{bmatrix} \begin{bmatrix} 20.16\% \\ 27.93\% \\ 22.21\% \\ 2.33\% \\ 12.58\% \\ 12.58\% \\ 1.11\% \\ 1.11\% \end{bmatrix}$$

Asset Class	Market Capitalization Estimate	Market Capitalization Weights w_{mkt}
US Bonds	$8,360,741,000,000	20.16%
Global Bonds xUSD	$11,583,275,710,000	27.93%
World Equity xUS	$9,212,460,000,000	22.21%
Emerging Equity	$964,647,000,000	2.33%
US Large Cap Growth	$5,217,844,438,500	12.58%
US Large Cap Value	$5,217,844,438,500	12.58%
US Small Cap Growth	$459,897,061,500	1.11%
US Small Cap Value	$459,897,061,500	1.11%
Total	$41,476,606,710,000	100.00%

Asset Class	Implied Excess Return		Risk-Free Rate		Total Implied Return
US Bonds	0.08%	+	4.00%	=	4.08%
Global Bonds xUSD	0.95%	+	4.00%	=	4.95%
World Equity xUS	3.95%	+	4.00%	=	7.95%
Emerging Equity	5.37%	+	4.00%	=	9.37%
US Large Cap Growth	5.13%	+	4.00%	=	9.13%
US Large Cap Value	3.68%	+	4.00%	=	7.68%
US Small Cap Growth	6.12%	+	4.00%	=	10.12%
US Small Cap Value	3.50%	+	4.00%	=	7.50%

El Modelo de Black-Litterman

La fórmula de Black-Litterman nos permite obtener un nuevo estimador para el vector de retornos:

$$E[R] = \left[(\tau\Sigma)^{-1} + P'\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q\right]$$

Donde:

$E[R]$	is the new (posterior) Combined Return Vector (*N x 1* column vector);
τ	is a scalar;
Σ	is the covariance matrix of excess returns (*N x N* matrix);
P	is a matrix that identifies the assets involved in the views (*K x N* matrix or *1 x N* row vector in the special case of 1 view);
Ω	is a diagonal covariance matrix of error terms from the expressed views representing the uncertainty in each view (*K x K* matrix);
Π	is the Implied Equilibrium Return Vector (*N x 1* column vector); and,
Q	is the View Vector (*K x 1* column vector).

Conclusiones

El análisis de media-varianza descansa sobre el supuesto de que inversionistas racionales seleccionan eligen activos riesgosos tomando en cuenta el retorno esperado del portafolio y el riesgo (medido como la varianza del portafolio)

Sin embargo, la reputación de este marco teórico clásico entre los *practitioners* se ha visto perjudicada debido a diversas dificultades en la implementación.

Los pesos de portafolio derivados son bastante sensibles a los inputs (ver Best & Grauer,1991), especialmente los retornos esperados, y a menudo representan asignaciones no intuitivas o extremas que exponen a un inversor a riesgos no intencionados.