Pronóstico de series de tiempo y simulaciones en R

Clase 4: Modelos con heteroscedasticidad condicional

Germán Eduardo González

Índice

Introducción y motivación

Modelo ARCH

Modelo GARCH

Extensiones del modelo GARCH

Aplicación

Índice

Introducción y motivación

Modelo ARCH

Modelo GARCH

Extensiones del modelo GARCH

Aplicación

Paréntesis pequeño

Teorema de representación de Wold

- Todo proceso estacionario de covarianza con media 0 se puede reescribir como MA infinito

$$Y_t = \sum_{i=1}^{\infty} b_i \epsilon_{t-i}$$

$$\epsilon_t \sim N(0, \sigma^2); \mathrm{b}_0 = 1; \sum_{i=1}^{\infty} b_i < \infty$$

Paréntesis pequeño

- Si este teorema es cierto, ¿Por qué usar ARIMAs y no solo MAs?
- De nuevo es por una razón de parsimonia.
- Son pocos los valores muy rezagados los que afectan el valor de hoy, condicional a los rezagos recientes.
- En ese orden de ideas **los rezagos AR reducen la cantidad de rezagos MA (plausiblemente grande) que debo incluir para tener un buen modelo.**

Introducción y motivación

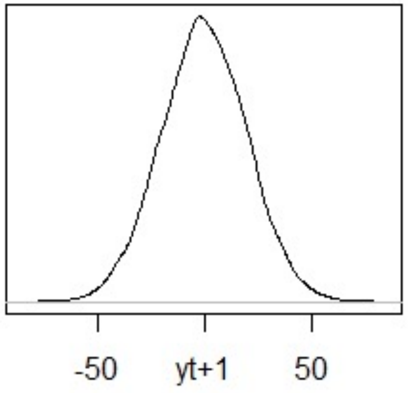
- En la sección anterior estudiamos especificaciones autorregresivas tipo ARIMA.
- Este tipo de especificaciones permite un buen pronostico punto **(promedio)** de la variable de interés.
- No obstante, la media no es la única variable de interés en muchos contextos.
  - Valoración de derivados (BSM por ejemplo).
  - Cálculo de riesgo.
  - Simulación de rendimientos de un portafolio.

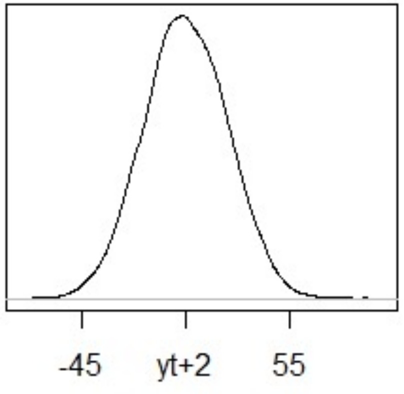
Introducción y motivación

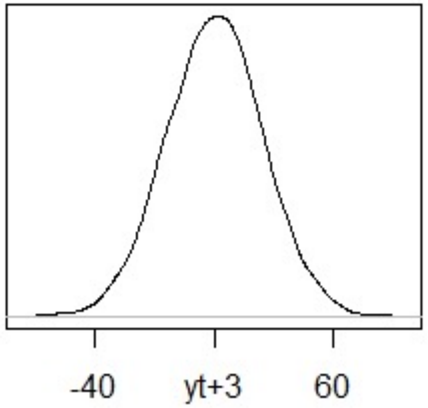
- Hasta el momento hemos estudiado este comportamiento:

$$Y_t = \beta_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma^2)$$





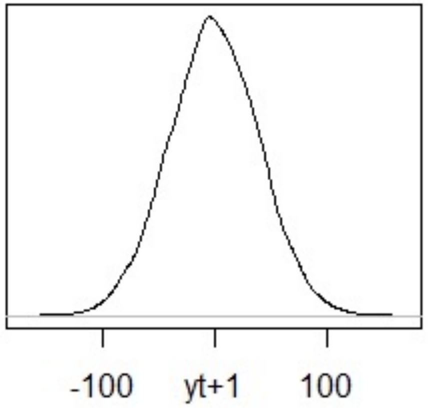


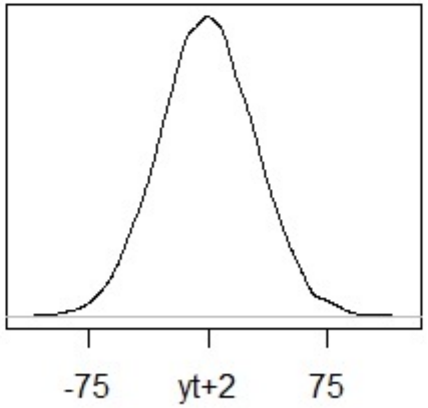
Introducción y motivación

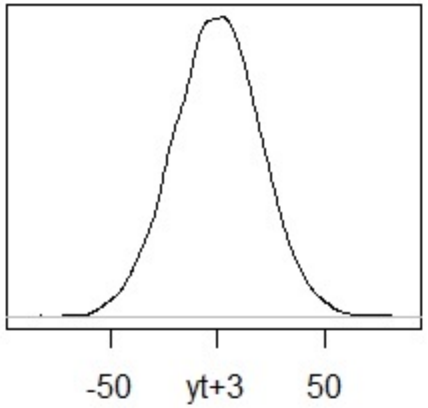
- Pero gran parte de las series financieras tienen uno parecido a este:

$$Y_t = \beta_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

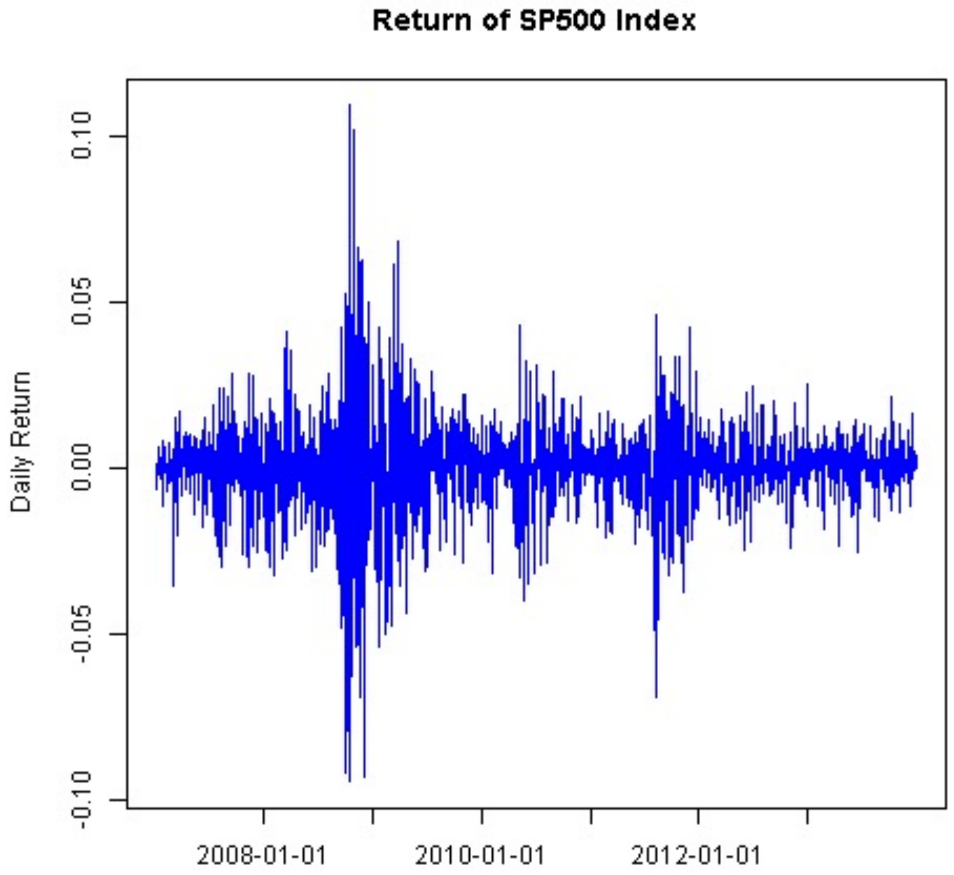
$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$







Introducción y motivación



Introducción y motivación

- Las series financieras suelen tener **clusters de volatilidad.**
- Lo que quiere decir que si los días anteriores fueron muy volátiles, lo más probable es que hoy lo sea también. Si los días anteriores fueron poco volátiles, hoy probablemente el comportamiento será parecido.
- Dicho de otra forma, la volatilidad de estas series parece ser **autorregresiva** en sí misma.

Índice

Introducción y motivación

Modelo ARCH

Modelo GARCH

Extensiones del modelo GARCH

Aplicación

Modelo ARCH

- La intuición previamente descrita fue la que llevo a Engle (1982) a proponer los modelos ARCH: Autoregresive Conditional Heteroskedasticity.
- Un shock importante en la variable en periodos anteriores lleva a volatilidades altas hoy **(Varianza condicional autorregresiva).**
- Pero cuando t tiende a infinito, o dicho de otra manera, cuando tengo muchos periodos, la varianza debería ser algo relativamente constante **(Varianza incondicional constante)**.

Modelo ARCH

- Recordando el teorema de representación de Wold **(todo proceso estacionario de covarianza con media 0 se puede reescribir como MA infinito),** teníamos antes que:

$$Y_t = \sum_{i=1}^{\infty} b_i \epsilon_{t-i}$$

$$\epsilon_t \sim N(0, \sigma^2); \mathrm{b}_0 = 1; \sum_{i=1}^{\infty} b_i < \infty$$

$$Var(Y_t) = \sigma^2 \sum_{i=1}^{\infty} b_i^2 \, ; Var(Y_t | \Omega_{t-1}) = \sigma^2$$

Modelo ARCH

- Y ahora:

$$Y_t = \sum_{i=1}^{\infty} b_i \epsilon_{t-i}$$

$$\epsilon_t | \Omega_{t-1} = N(0, \sigma_t)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \gamma_i \epsilon_{t-i}^2$$

$$Var(Y_t) = \frac{\omega}{1 - \sum_{i=1}^{p} \gamma_i} \sum_{i=1}^{\infty} b_i^2 \,; Var(Y_t | \Omega_{t-1}) = \omega + \sum_{i=1}^{p} \gamma_i \epsilon_{t-i}^2$$

Modelo ARCH

- Incorpora el efecto de los choques en distancia a la media de los periodos pasados, incondicional a su efecto a través de la volatilidad misma de los periodos pasados **(Análogo al MA en un ARIMA).**
- Con esta intuición en mente, una forma clásica de definir los rezagos del ARCH(P) es con autocorrelogramas **(Igual que el MA).**

Índice

Introducción y motivación

Modelo ARCH

Modelo GARCH

Extensiones del modelo GARCH

Aplicación

Modelo GARCH

- Continuando con la analogía con el MA, si bien se comporta como un modelo sencillo para estimar, el ARCH suele ser poco parsimonioso al necesitar de muchos rezagos para pronosticar correctamente la volatilidad.
- Sin embargo, **(análogo al componente AR)** condicional al efecto que ya tuvo en la volatilidad de periodos anteriores, no hay muchos rezagos que tengan un efecto directo.

Modelo GARCH

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \gamma_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

Y así llegamos al GARCH(P,Q): Generalized Autoregresive Conditional Heteroskedasticity propuesto por Bollerslev (1986).

Modelo GARCH

- En este orden de ideas, parece intuitivo suponer que así como los rezagos Q se pueden escoger con ACF, los rezagos P se escogen con el PACF.
- Sin embargo, algunas investigaciones se han hecho (Miah, et al (2016), Namugaya, et al (2014), Bahraminasab, et al (2007)) para justificar que, no solo por parsimonia sino por desempeño, GARCH(1,1) es suficiente en la práctica para capturar dinámicas autorregresivas en la volatilidad, sobre todo de series financieras.

Modelo GARCH

- Charles, W., Naguyama, J. & Weke, P. (2014) Modelling Volatility of Stock Returns: Is GARCH(1,1) Enough?. Recuperado de: https://core.ac.uk/download/pdf/249334038.pdf
- Mia, M., Rahman, A. (2016). Modelling Volatility of Daily Stock Returns: Is GARCH(1,1) Enough? Recuperado de: https://www.researchgate.net/publication/335749475_Modelling_Volatility_of_Daily_Stock_Returns_Is_GARCH11_Enough
- Bahraminasab, A., Jafari, G. Norouzzadeh, P. (2007). Why does the standard GARCH(1, 1) model work well? Recuperado de: https://www.researchgate.net/publication/242308122_Why_does_the_standard_GARCH1_1_model_work_well

Índice

Introducción y motivación

Modelo ARCH

Modelo GARCH

Extensiones del modelo GARCH

Aplicación

Extensiones del modelo GARCH

- Hay algunas consideraciones que se dejan por fuera en la especificación básica del GARCH.
  - ¿La volatilidad puede afectar la media?
  - ¿Hay variables exógenas que puedan afectar la volatilidad?
  - ¿Cambios de volatilidad por región?

GARCH M

- Garch in mean.
- Es la única especificación hasta el momento que cambia también el pronóstico punto.
- Incorpora la noción de que la alta volatilidad puede cambiar la media condicional.
- Esto suele suceder en activos de alta bursatilidad, momentos de alta volatilidad pueden llevar a un disparo del precio o a una caída del mismo.

$$Y_t = \beta_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \psi \sigma_t^2 + \epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \gamma \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

GARCH con variables exógenas

- Permite incluir variables que puedan ser predictivas de la volatilidad (por ejemplo el volumen del mercado).

$$\sigma_t^2 = \omega + \gamma\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \phi X_t$$

Uniendo varias especificaciones

- En el caso de la TRM, uno puede estar convencido en que los retornos de esta están fuertemente afectados por los del Brent.
- Este efecto puede ser a través del efecto que tiene el Brent en la volatilidad de la TRM y como esta tiene efecto en la media, y también un efecto directo.
- Quedaría un GARCH M con variables exógenas.

$$Y_t = \beta_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \psi \sigma_t^2 + \alpha X_t + \epsilon_t$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \gamma \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \phi X_t$$

T GARCH

- Threshold GARCH.
- Permite que el efecto marginal en la volatilidad difiera para rangos de ϵ_t.

$$\sigma_t^2 = \omega + \gamma\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \psi\epsilon_{t-1}^2 \times D_i$$

$$D_i = \begin{cases} 1, & \epsilon_t < a \\ 0, & dlc \end{cases}$$

- En la práctica (e intuitivamente a menos que se tuviera un supuesto especial ex-ante) $a = 0$.

Índice

Introducción y motivación

Modelo ARCH

Modelo GARCH

Extensiones del modelo GARCH

Aplicación

Frontera Eficiente

- Una frontera eficiente consiste en el conjunto de portafolios que minimizan alguna medida de riesgo dado una rentabilidad mínima.

- El primero en pensar en este concepto e implementarlo fue Markowitz (1952).

- El principio, aunque simple, es extremadamente fuerte, ya que permite capturar a todos los mejores portafolios sin necesidad de generar infinitas combinaciones.

Correlación de los Retornos (12M): 27/06/2019

	COLCAP	Ecopetrol	Bancolombia PF	SURA PF	Nutresa	ISA	Aval PF	Cemargos	GEB	Argos PF	Davivienda PF	Éxito	Bogota	Celsia	Cemex
Cemex	−41%	−19%	−24%	−48%	−10%	−37%	−9%	−25%	0%	−11%	−23%	−20%	0%	−16%	
Celsia	−11%	0%	−17%	−24%	−15%	0%	1%	−17%	−10%	8%	−9%	4%	−15%		−16%
Bogota	−11%	−6%	−8%	−5%	12%	−25%	12%	−28%	−19%	13%	−7%	−21%		−15%	0%
Éxito	−14%	−24%	−21%	−6%	−7%	15%	20%	22%	−1%	−27%	16%		−21%	4%	−20%
Davivienda PF	16%	−7%	16%	13%	4%	1%	−5%	20%	7%	4%		16%	−7%	−9%	−23%
Argos PF	10%	−10%	−2%	34%	32%	−9%	−31%	6%	8%		4%	−27%	13%	8%	−11%
GEB	−7%	−6%	−21%	1%	31%	−10%	−24%	−7%		8%	7%	−1%	−19%	−10%	0%
Cemargos	26%	−1%	16%	37%	−15%	19%	3%		−7%	6%	20%	22%	−28%	−17%	−25%
Aval PF	27%	7%	25%	5%	−33%	27%		3%	−24%	−31%	−5%	20%	12%	1%	−9%
ISA	31%	−6%	13%	40%	14%		27%	19%	−10%	−9%	1%	15%	−25%	0%	−37%
Nutresa	3%	−16%	−7%	20%		14%	−33%	−15%	31%	32%	4%	−7%	12%	−15%	−10%
SURA PF	67%	25%	47%		20%	40%	5%	37%	1%	34%	13%	−6%	−5%	−24%	−48%
Bancolombia PF	83%	51%		47%	−7%	13%	25%	16%	−21%	−2%	16%	−21%	−8%	−17%	−24%
Ecopetrol	74%		51%	25%	−16%	−6%	7%	−1%	−6%	−10%	−7%	−24%	−6%	0%	−19%
COLCAP		74%	83%	67%	3%	31%	27%	26%	−7%	10%	16%	−14%	−11%	−11%	−41%

Frontera Eficiente

- Algunos supuestos:
  - El riesgo del portafolio solo está dado por la variabilidad de los resultados (la varianza)
  - Los individuos siempre prefieren **mayor retorno** y **menor riesgo**.
  - Existe un activo libre de riesgo, sin incertidumbre.
  - Hay divisibilidad perfecta del mercado.
  - Los **precios** son **perfectamente observables** y las matrices de varianza-covarianza, así como los retornos esperados son conocidos.
  - Es posible irse en corto (aunque no lo veremos a profundidad en esta clase).
  - Existen solo dos periodos: 0 y 1.

Frontera Eficiente

En cualquier momento del tiempo, el valor del portafolio es:

$$V(t) = \sum_i x_i * S_i(t)$$

Donde:

- x_i es el número de acciones *i* en el portafolio
- $S_i(t)$ es precio de la acción *i*
- $V(t)$ es el valor del portafolio en el tiempo t.

En adelante, definamos el peso de un activo cualquiera en el portafolio en el tiempo t como:

$$w_{i,t} = \frac{x_i * S_i(t)}{V(t)}$$

Al ser el valor del portafolio la suma del valor de todos los activos en este:

$$\sum_i w_i = 1$$

Caso 2 activos

Dónde el valor en T=1 es igual a"

$$S_1(1) = S_1(0) * (1 + r_1)$$

$$S_2(1) = S_2(0) * (1 + r_2)$$

- Para el caso de dos activos, calculemos el retorno esperado del portafolio:

$$E(r) = E\left(\frac{V(1)}{V(0)}\right) - 1 = E\left(\frac{x_1 * S_1(1) + x_2 * S_2(1)}{V(0)}\right) - 1$$

"Desfactorizando"

$$= E\left(\frac{x_1 * S_1(0) * (1 + r_1) + x_2 * S_2(0) * (1 + r_2)}{V(0)}\right) - 1$$

$$E\left(\frac{\left(x_1 * S_1(0) + x_2 * S_2(0)\right) + x_1 * S_1(0) * r_1 + x_2 * S_2(0) * r_2}{V(0)}\right) - 1$$

$$E\left(\frac{V(0) + (x_1 * S_1(0) * r_1 + x_2 * S_2(0) * r_2)}{V(0)}\right) - 1 = E\left(1 + \frac{x_1 * S_1(0) * r_1 + x_2 * S_2(0) * r_2}{V(0)}\right) - 1$$

Usando la definición de pesos

$$w_{i,t} = \frac{x_i * S_i(t)}{V(t)}$$

$$E(r) = E(r_1 * w_1 + r_2 w_2) = w_1 * \mu_1 + w_2 * \mu_2$$

$$r_p = w_1 * r_1 + w_2 * r_2$$

Caso 2 activos

- Ahora miremos la varianza

$$V(r) = V(r_1 * w_1 + r_2 w_2) = V(r_1 * w_1) + V(r_2 * w_2) + 2Cov(r_1 * w_1 + r_2 w_2)$$
$$V(r) = w_1^2 * V(r_1) + w_2^2 * V(r_2) + 2 * w_1 * w_2 * Cov(r_1, r_2)$$

- Definiciones para simplificar:

$$\sigma_1 = \sqrt{V(r_1)}; \sigma_2 = \sqrt{V(r_2)}\ ;\ \rho_{1,2} = \frac{Cov(r_1, r_2)}{\sigma_1 * \sigma_2}$$

- Entonces podemos reescribir la varianza como:

$$V(r) = w_1^2 * \sigma_1^2 + w_2^2 * \sigma_2^2 + 2 * w_1 * w_2 * \rho_{1,2} * \sigma_1 * \sigma_2$$

Caso 2 activos

Teorema

- Se pueden encontrar **todos los puntos de una frontera eficiente (**media varianza) haciendo todas las combinaciones **convexas** de dos portafolios dentro de esta.
- Cuando se tienen dos activos posibles nada más, ambas soluciones de esquina se encuentran dentro de la frontera.
- Entonces las combinaciones lineales entre los dos mapean la frontera.
- Este teorema deja de ser del todo cierto cuando se tienen restricciones.
- En cursos más avanzados se pueden ver excepciones a este teorema, para el ejercicio práctico de esta clase usaremos un caso en el que se cumple.

Caso 2 activos

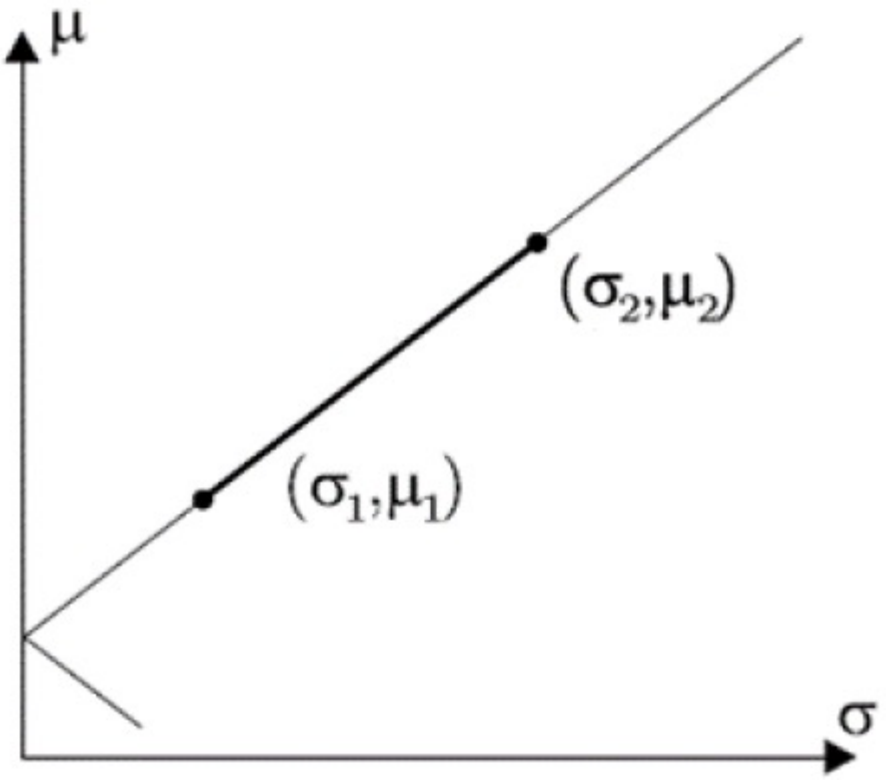
- Si los activos están perfectamente correlacionados positivamente:
  - $\rho_{1,2} = 1$

$$V(r) = w_1^2 * \sigma^2 + w_2^2 * \sigma^2 + 2 * w_1 * w_2 * \sigma_1 * \sigma_2$$
$$= (w_1 * \sigma_1 + w_2 * \sigma_2)^2$$
$$\sigma = |w_1 * \sigma_1 + w_2 * \sigma_2|$$

- Por último, usemos la restricción de pesos del portafolio:

$$\sigma = |w_1 * \sigma_1 + (1 - w_1) * \sigma_2|$$

- Solo se puede reducir el riesgo tomando posiciones cortas en activos riesgosos.



Caso 2 activos

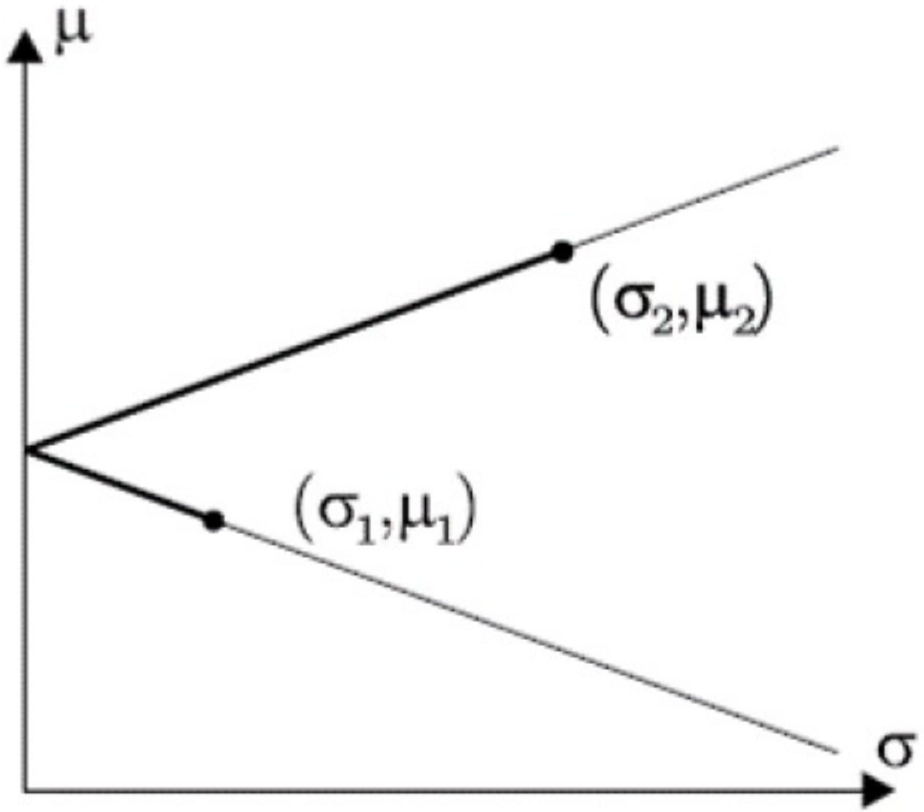
- Si los activos están perfectamente correlacionados negativamente:
  - $\rho_{1,2} = -1$

$$V(r) = w_1^2 * \sigma^2 + w_2^2 * \sigma^2 - 2 * w_1 * w_2 * \sigma_1 * \sigma_2$$
$$= (w_1 * \sigma_1 - w_2 * \sigma_2)^2$$
$$\sigma = |w_1 * \sigma_1 - w_2 * \sigma_2|$$

- Por último, usemos la restricción de pesos del portafolio:

$$\sigma = |w_1 * \sigma_1 - (1 - w_1) * \sigma_2|$$

- El riesgo se reduce con una combinación de ambos activos.



Caso 2 activos

- Por último, encontremos el punto de mínima varianza:

$$\min_{w_1} w_1^2 * \sigma_1^2 + (1-w_1)^2 * \sigma_2^2 + 2 * w_1 * (1-w_1) * \rho_{1,2} * \sigma_1 * \sigma_2$$

$$\frac{\partial V(r)}{\partial w_1} = 2w_1 * \sigma_1^2 - 2(1-w_1) * \sigma_2^2 + 2 * (1-2w_1)\rho_{1,2} * \sigma_1 * \sigma_2 = 0$$

$$2w_1 * \sigma_1^2 + 2w_1 * \sigma_2^2 - 4w_1\rho_{1,2} * \sigma_1 * \sigma_2 = 2\sigma_2^2 - 2\rho_{1,2} * \sigma_1 * \sigma_2$$

$$w_1(\sigma_1^2 + \sigma_2^2 - 2\rho_{1,2} * \sigma_1 * \sigma_2) = \sigma_2^2 - \rho_{1,2} * \sigma_1 * \sigma_2$$

$$w_1 = \frac{\sigma_2^2 - \rho_{1,2} * \sigma_1 * \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho_{1,2} * \sigma_1 * \sigma_2}$$

Caso N activos

- Ahora analicemos las diferencias con N activos.
- Existe una matriz de Varianza-Covarianza (dada):

$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \cdots & \sigma_{1,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{N,1} & \sigma_{1,1} & \cdots & \cdots & \sigma_{N,N} \end{bmatrix}; donde\ \sigma_{i,i} = \sigma_i^2$$

Supondremos que la matriz Σ es invertible. Solo se incumple si un activo es replicable por una combinación de otros activos o si un activo tiene varianza cero.

- Además, existe un vector de retornos esperados:

$$\mathrm{M} = [\mu_1, \mu_2, \dots, \mu_N]$$

- El retorno del portafolio se descompone en el retorno de los activos que lo componen:

$$r_p = w_1 * r_1 + \cdots + w_N * r_N = \sum_i w_i * r_i$$

Caso N activos

- Utilizando el resultado anterior:

$$E(r_p) = \mu_p = \sum_i w_i * E(r_i) = \sum_i w_i * \mu_i$$

- De manera matricial:

$$\mu_p = \mathrm{M} * w^T$$

- La varianza:

$$V(r_p) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i * w_j * \sigma_{i,j} = w * \Sigma * w^T$$

Caso N activos

- Para este nuevo caso de N activos, la frontera no surge de la combinación de dos activos, sino de la utilización al mismo tiempo de distintos pesos para distintos activos. **En pocas palabras, no tiene forma cerrada.** De esta manera, definimos un algoritmo para generar dicha frontera:

1. Definimos un conjunto de retornos objetivo.

2. Para cada retorno objetivo, encontramos el punto de mínima varianza.

3. El conjunto de puntos (portafolios) encontrados representa la frontera eficiente.

Afortunadamente, el problema del punto 2 es convexo y se puede encontrar a partir de derivadas (computacionalmente fácil):

Caso N activos

- Objetivo: Encontrar el portafolio de mínima varianza.

Recordemos que:

$$p * Var(x) + (1-p) * Var(y) \geq Var(p * x + (1-p) * y)$$

Lo anterior nos define un problema convexo, lo que permite encontrar relativamente la solución a problema complejos, (i) sin la necesidad de mapear todos los escenarios y (ii) con un tiempo de optimización considerablemente bajo.

El problema a resolver es:

$$\min_{w} w * \Sigma * w^T \; s.a. \; \mathrm{M} * w^T = e$$

Gracias

Universidad de los Andes

quantil

matemáticas aplicadas