Seminario Aleatorio

Un enfoque de regresión a selección de Portafolio

Jorge de la Vega

Banco de México

Resumen

- El objetivo de esta platica es *introducir*, con un ejemplo sencillo, cómo se puede plantear el problema de portafolio de mínima varianza como un problema de mínimos cuadrados y cómo se puede extender el problema en un contexto bayesiano que permite introducir dinámica en la estimación.
- La aplicación se basa en un trabajo mucho más extenso no publicado de Mark-Britten Jones (LSE).

Definición del problema

- Contexto:
  - Se cuenta con p activos, y observaciones de sus rendimientos en T periodos.
  - Sea $\mathbf{r}_t$ el vector de rendimientos en t.
- El portafolio global de varianza mínima (GVM) es un vector $\boldsymbol{\omega}_g$ de pesos que satisface el siguiente problema de optimización:

$$\boldsymbol{\omega}_g = arg \min_{\{\boldsymbol{\omega} | \boldsymbol{\omega}' \mathbf{1} = 1\}} \mathrm{Var}(\mathbf{r}_t' \boldsymbol{\omega})$$

Un problema de mínimos cuadrados

- El problema de optimización se puede replantear como un problema de regresión del siguiente modo: particionando los vectores de rendimientos y pesos:

$$\mathbf{r}_t = \begin{pmatrix} r_{t,1} \\ \mathbf{r}_{t,2} \end{pmatrix}, \boldsymbol{\omega} = \begin{pmatrix} w_1 \\ \mathbf{w} \end{pmatrix}$$

  entonces $\mathbf{r}_t \boldsymbol{\omega} = r_{t,1} + \mathbf{r}_t'^e \mathbf{w}$, donde $\mathbf{r}_t^e = \mathbf{r}_{t,2} - r_{t,1}\mathbf{1}$ es el vector de excesos de rendimientos con respecto al instrumento base (en este caso, el primero).

- Si escribimos el modelo de regresión con respuesta $y = -r_{t,1}$ y predictores $\mathbf{x} = \mathbf{r}_t^e$:

$$-r_{t,1} = -\alpha + \mathbf{r}_t'^e \mathbf{w} + u_t, \qquad u_t \sim \mathcal{N}\left(0, \sigma^2\right), \quad t = 1, 2, \ldots, T$$

  entonces $\sigma^2 = \mathrm{Var}(u_t) = \mathrm{Var}(\alpha - r_{t,1} - \mathbf{r}_t'^e \mathbf{w}) = \mathrm{Var}(r_{t,1} + \mathbf{r}_t'^e \mathbf{w}) = \mathrm{Var}(\mathbf{r}_t' \boldsymbol{\omega})$

- Así el problema de portafolio es equivalente a un problema de mínimos cuadrados (minimizar $\sum_{t=1}^{T} u_t^2$) sobre $\mathbf{w}$.

Simplificando el problema

- En términos matriciales, si $\mathbf{r}_1 = \begin{pmatrix} r_{1,1} \\ \vdots \\ r_{T,1} \end{pmatrix}$ y $\mathbf{R}_e = \begin{pmatrix} \mathbf{r}'_{1,2} \\ \vdots \\ \mathbf{r}'_{T,2} \end{pmatrix}$, el modelo se puede expresar como

$$-\mathbf{r}_1 = -\alpha\mathbf{1} + \mathbf{R}_e\mathbf{w} + \boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$$

- Como α es irrelevante en el contexto del portafolio, podemos quitarla del modelo centrando los datos con la matriz $\mathbf{M} = \mathbf{I} - \frac{1}{T}\mathbf{1}\mathbf{1}'$. (notar que $\mathbf{M}\mathbf{1} = \mathbf{0}$).

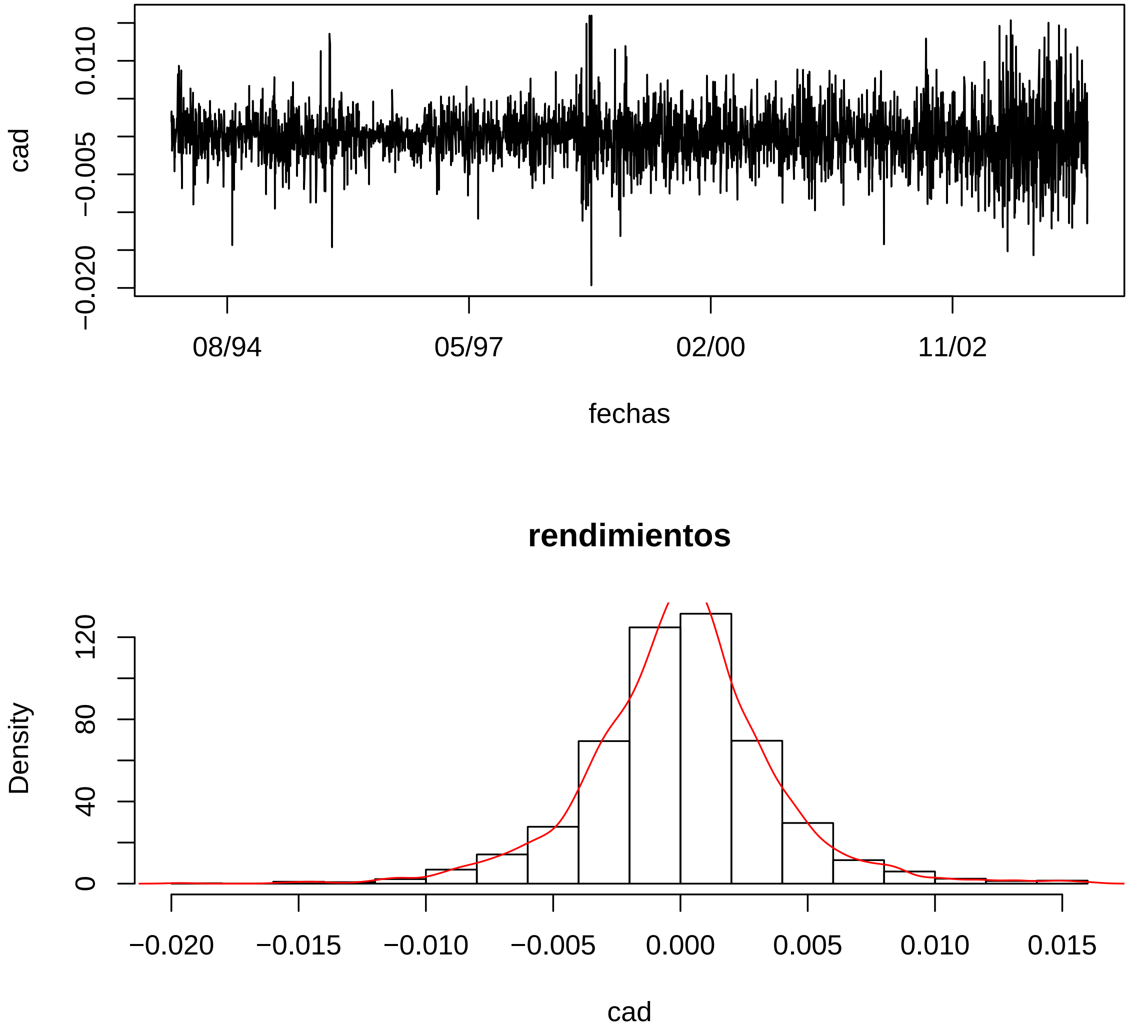
- Finalmente, haciendo $\mathbf{x}_1 = \mathbf{M}\mathbf{r}_1$, $\mathbf{X} = \mathbf{M}\mathbf{R}_e$ y $\mathbf{e} = \mathbf{M}\boldsymbol{\epsilon}$, el modelo queda como

$$-\mathbf{x}_1 = \mathbf{X}\mathbf{w} + \mathbf{e}, \qquad \mathbf{e} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{M}\right),$$

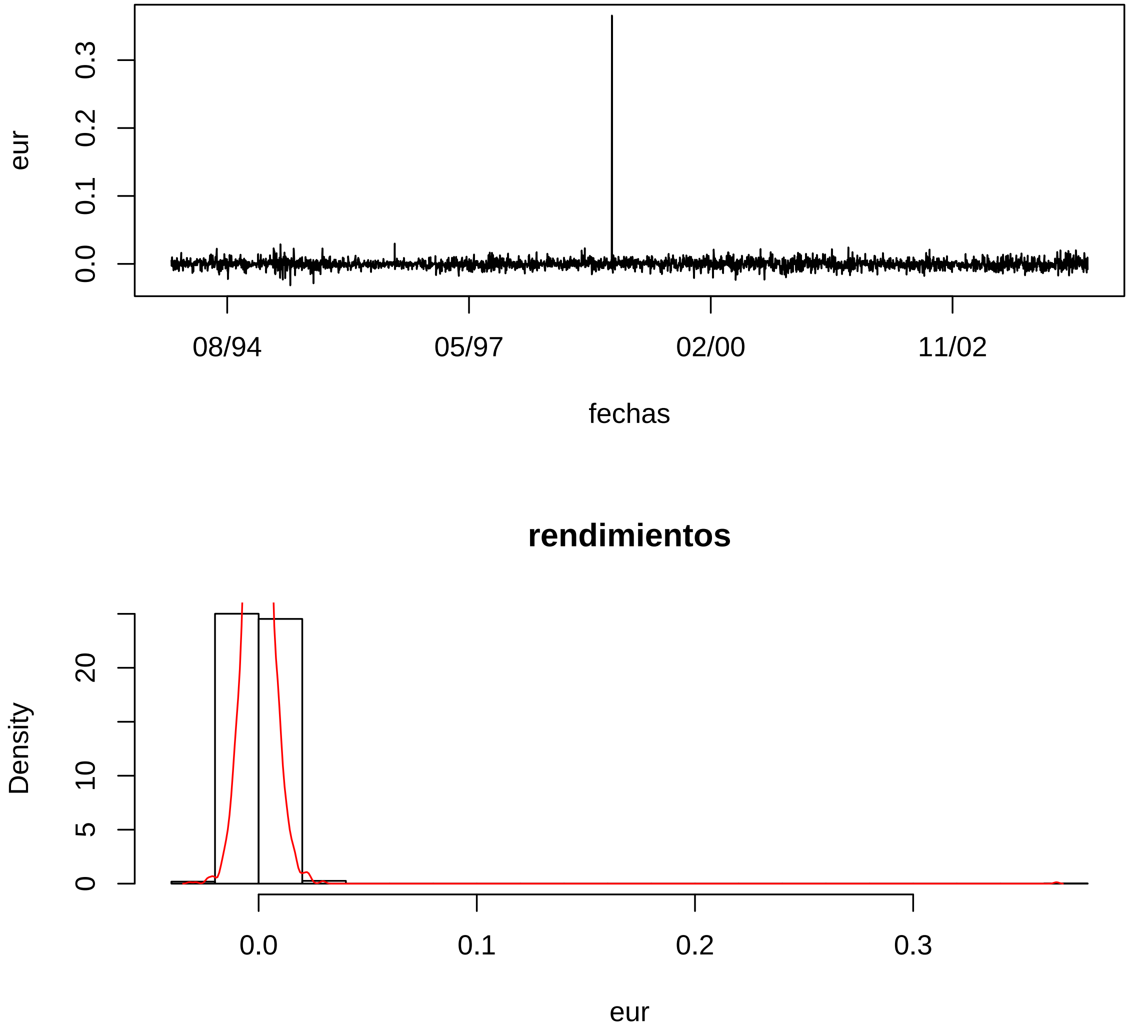
Ejemplo

- Los rendimientos considerados son los de los tipos de cambio de las divisas del BMK: cad, chf, eur, gbp y jpy, diarios. Todos los tipos de cambio está en divisa-usd. La ventana muestral abarca de 3 de Enero de 1994 al 21 de mayo del 2004.
- En el ejemplo se toma como divisa base al euro.
- Los rendimientos se toman diarios, pero pueden recalcularse en ventanas semanales, mensuales, etc. Esto acorta la muestra de rendimientos disponibles y por lo tanto la calidad de las estimaciones.
- En el caso del euro se tuvo que eliminar el rendimiento del 30/12/98, ya que es un valor bastante atípico (Entra en vigor el euro en enero 99).
- En el ejemplo, $T = 2708$, $p = 5$.

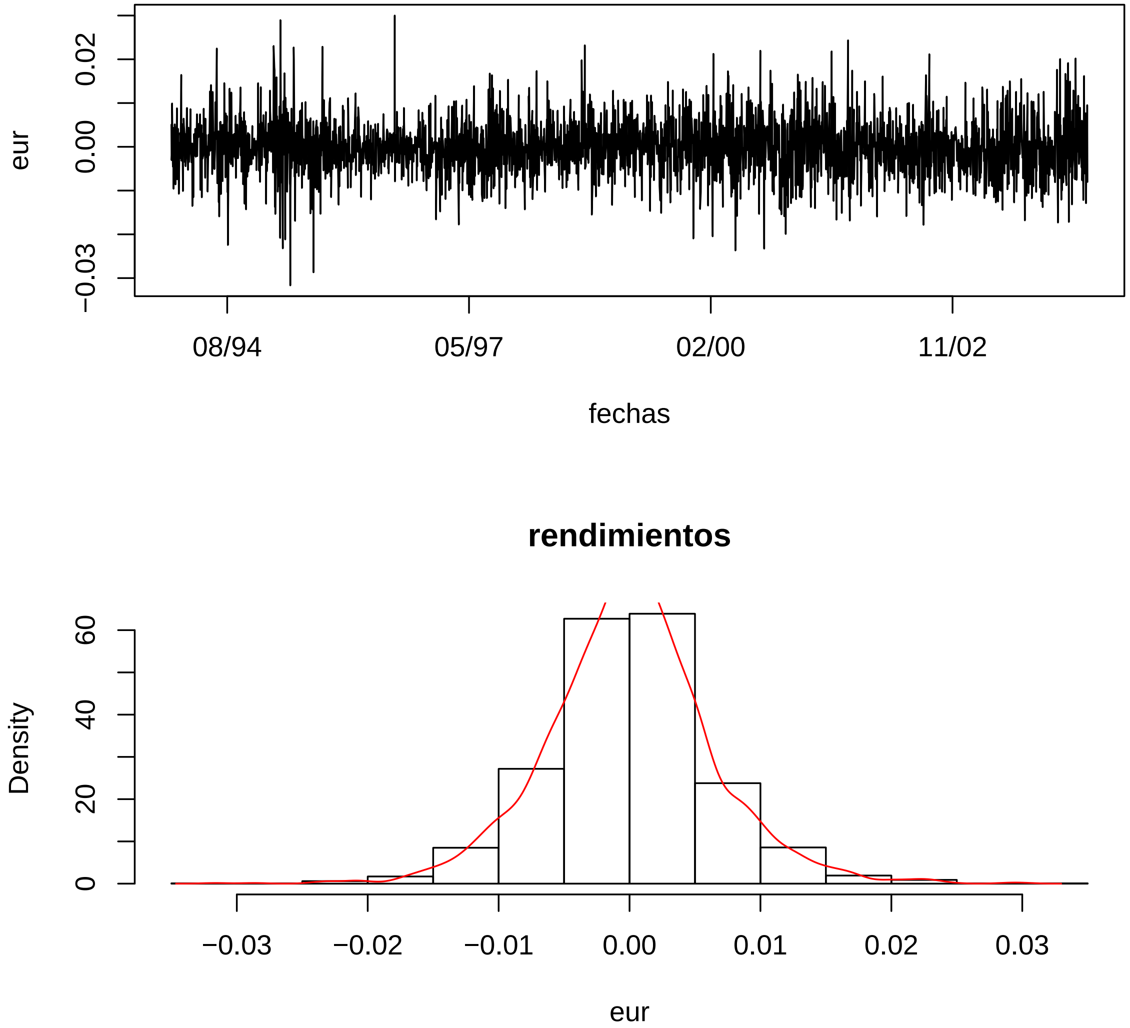
Distribución cad



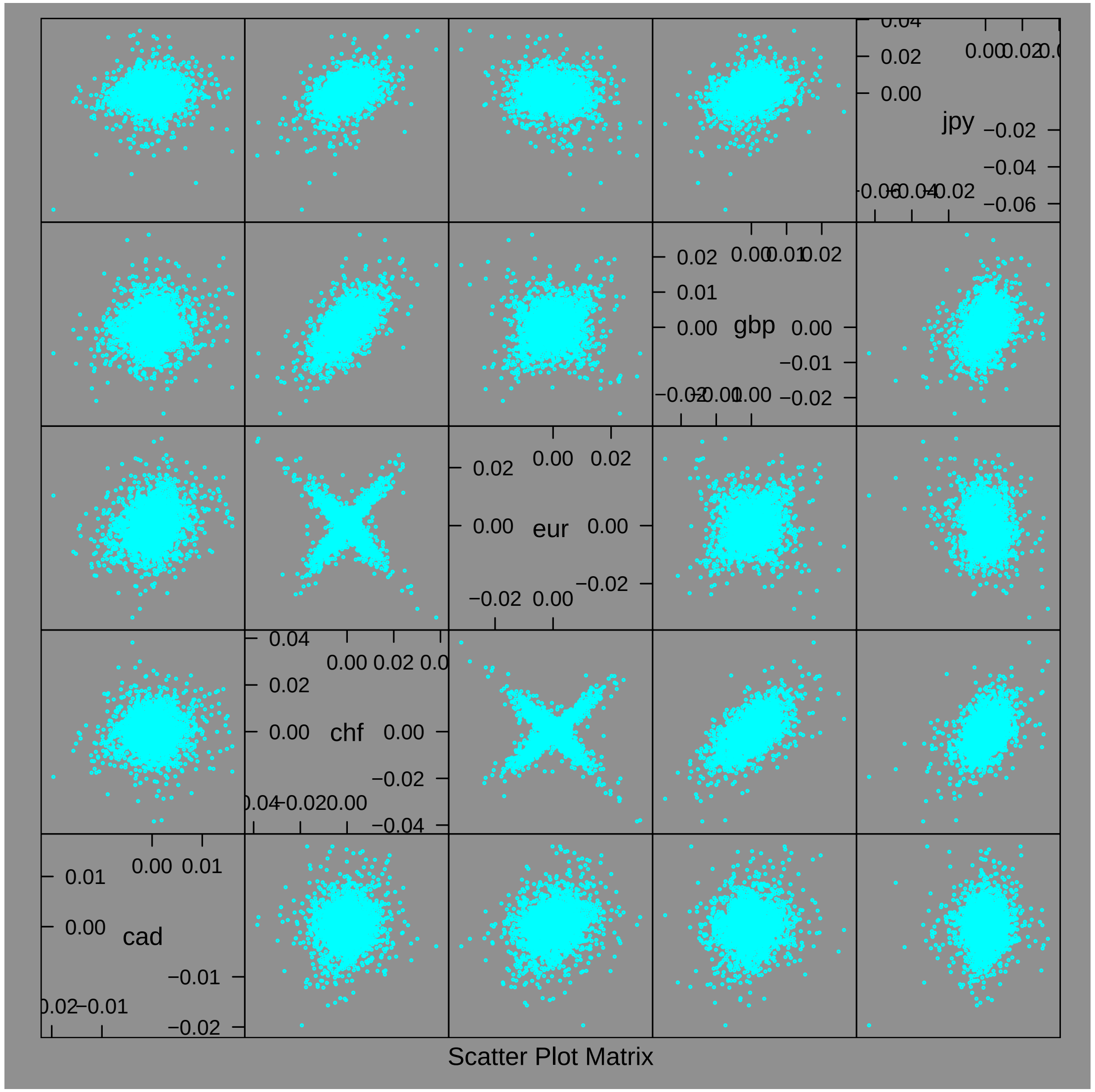
Distribución eur



Distribución eur



Correlación rendimientos



Estimación del portafolio

```
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
Xcad  0.521914   0.011972  43.596   <2e-16
Xchf -0.013272   0.010600  -1.252    0.211
Xgbp  0.241432   0.013271  18.193   <2e-16
Xjpy  0.111760   0.008234  13.573   <2e-16

Residual standard error: 0.002856 on 2704 degrees of freedom
Multiple R-Squared: 0.7949,     Adjusted R-squared: 0.7946
F-statistic:  2620 on 4 and 2704 DF,  p-value: < 2.2e-16
```

- El portafolio sugerido para el vector (eur,cad,chf,gbp,jpy) es entonces (en porcentajes): $\hat{\omega}_0 = (13.82, 52.19, -1.33, 24.14, 11.18)$.
- El coeficiente negativo para chf no es significativo, por lo que podemos calibrar el modelo para tener ponderaciones adecuadas: $\hat{\omega} = (13.74, 52.25, 0, 23.19, 10.82)$.

Observaciones

- El error estándar de los rendimientos se puede estimar *antes* de centrar los datos: $\hat{\sigma} = 0.002856$ o 28.56 pb.
- No hay garantía en general de que las posiciones cortas sean no significativas, por lo que es necesario buscar un modelo que incorpore la restricción de que sólo las posiciones largas sean admisibles.
- Una ventaja adicional del modelo de regresión es que permite no sólo valores puntuales, sino la posibilidad de considerar regiones de confianza para los pesos, lo que permite considerar una región de portafolios en la que los inversionistas pueden considerar su elección de cartera óptima.

Introduciendo un enfoque bayesiano

- El modelo anterior se puede escribir en un contexto más bayesiano, condicionando en los parámetros:

$$-\mathbf{x}_1|(\mathbf{X},\mathbf{w},\sigma^2) \sim \mathcal{N}\left(\mathbf{X}\mathbf{w},\sigma^2\mathbf{M}\right)$$

- Un modelo no es bayesiano sin distribuciones iniciales y posteriores de los parámetros.
- Hay dos posibilidades a considerar con respecto a los parámetros:
  i. Suponer σ^2 conocida.
  ii. Suponer σ^2 desconocida e incorporar un componente para su distribución apriori. Este supuesto complica considerablemente la estimación del modelo.
- Los predictores siempre se suponen conocidos.

Caso 1: σ^2 conocida

- Podemos imponer una prior sobre $\mathbf{w}$ y crear un modelo jerárquico:

$$\begin{aligned} -\mathbf{x}_1|\mathbf{X}, \mathbf{w} &\sim \mathcal{N}\left(\mathbf{X}\mathbf{w}, \sigma^2\mathbf{M}\right) \\ \mathbf{w} &\sim F(\mathbf{m}_0) \end{aligned}$$

- Sin información específica sobre los factores que determinan los rendimientos de los activos que forman el portafolio, se pueden imponer restricciones de carácter estadístico para obtener una prior adecuada.
  - Por simplificación computacional, se puede suponer una prior normal multivariada:

$$\mathbf{w} \sim \mathcal{N}\left(\mathbf{w}_0, \mathbf{C}_0\right)$$

  - La prior debe ser intercambiable: la distribución conjunta de los pesos no depende de la ordenación de los activos,pero hay suficiente información inicial para distinguir entre activos. Esta restricción determina los parámetros iniciales: $\mathbf{w}_0 = \frac{1}{n}\mathbf{1}$ y $\mathbf{C}_0 = a^{-1}\mathbf{E}$.

Caso 1: σ^2 conocida

- $\mathbf{E}$ es una matriz de correlaciones de dimensión $p-1$ la forma

$$\mathbf{E} = \begin{pmatrix} 1 & -\frac{1}{p-1} & -\frac{1}{p-1} & \cdots & -\frac{1}{p-1} \\ -\frac{1}{p-1} & 1 & -\frac{1}{p-1} & \cdots & -\frac{1}{p-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\frac{1}{p-1} & -\frac{1}{p-1} & -\frac{1}{p-1} & \cdots & 1 \end{pmatrix}$$

- Con esta información, la distribución posterior es otra normal multivariada $\mathcal{N}\left(\mathbf{m}, \mathbf{V}\right)$ con

$$\mathbf{m} = \mathbf{V}(\sigma^2\mathbf{X}'\mathbf{X}\hat{\mathbf{w}} + ap^{-1}\mathrm{E}^{-1}\mathbf{1})$$

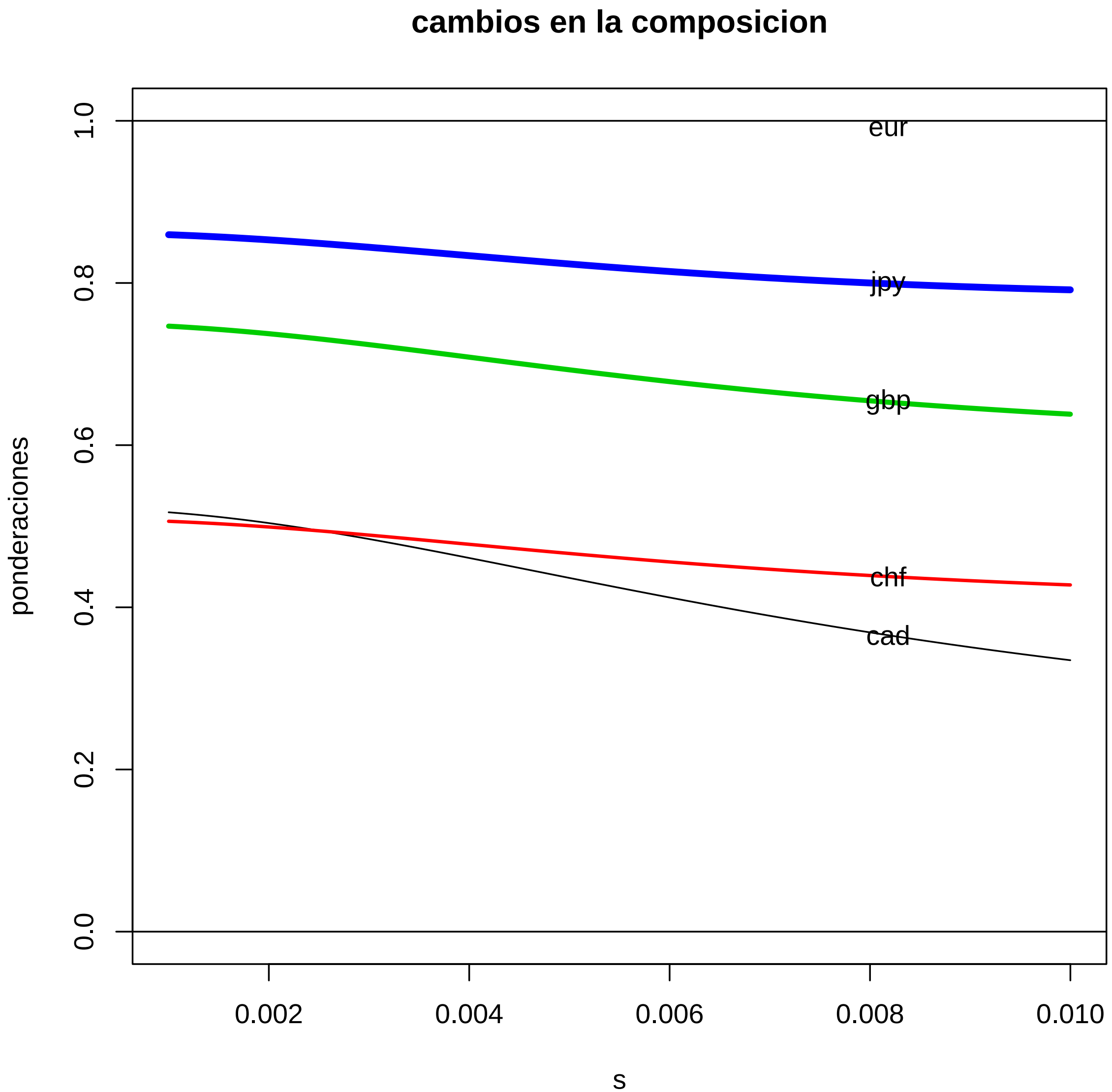
donde $\hat{\mathbf{w}} = -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{x}_1$, y

$$\mathbf{V} = (\sigma^2\mathbf{X}'\mathbf{X} + a\mathbf{E}^{-1})^{-1}$$

- Ahora el problema sólo depende de un parámetro desconocido, que es a.
- Resultado: La media de la distribución posterior es el conjunto óptimo de pesos que minimizan la varianza del rendimiento del portafolio.

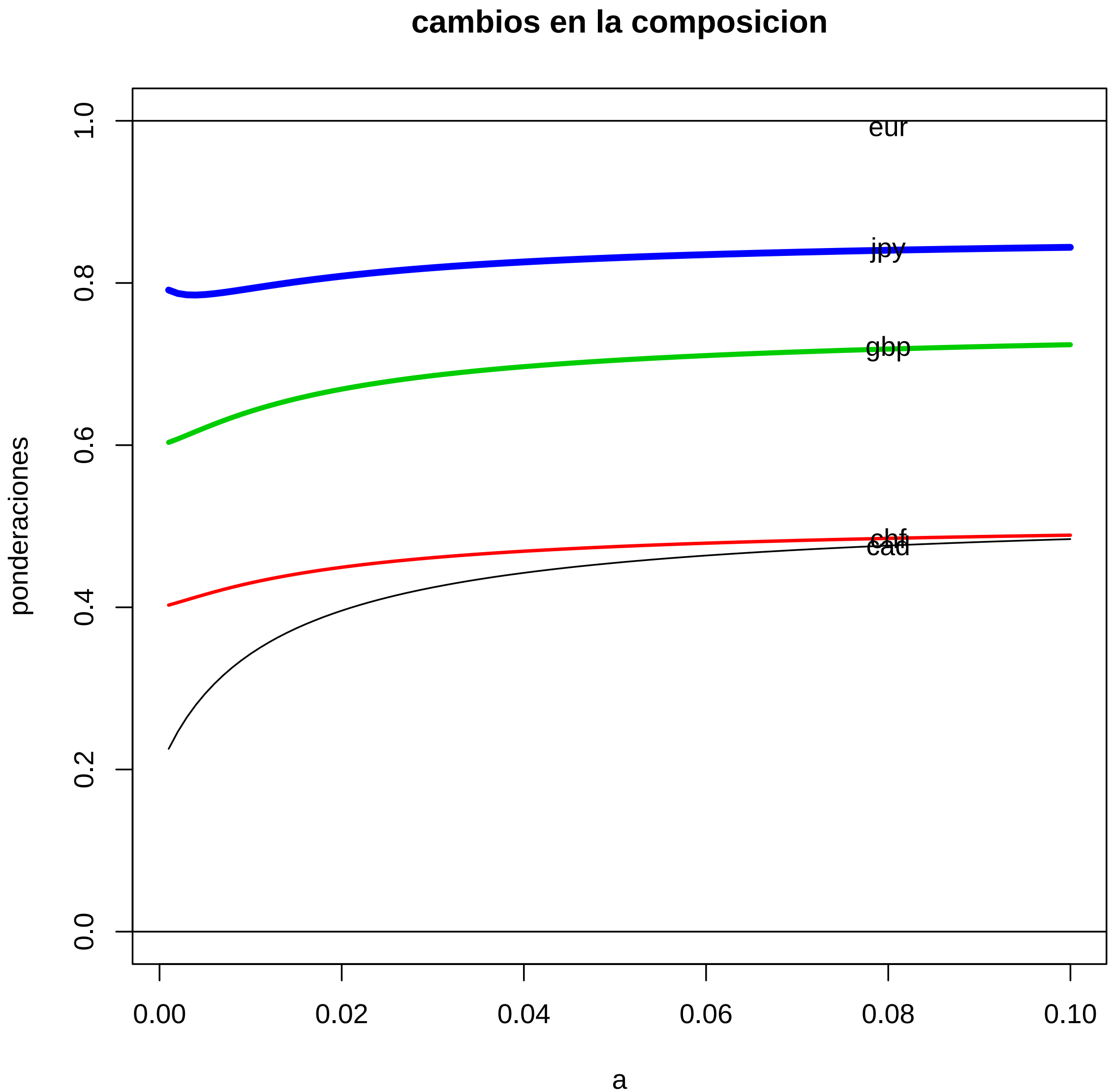
Ejemplo

- La siguiente gráfica muestra la composición para diferentes valores de σ y con un valor fijo de $a = 0.001$.



Ejemplo

- La siguiente gráfica muestra la composición para diferentes valores de a y con un valor fijo de $\sigma = 0.03$.



Versión dinámica del enfoque Bayesiano

- Ahora consideramos un modelo en donde los coeficientes pueden variar en el tiempo:

$$-r_{t,1} = -\alpha_t + \mathbf{r}_t^{e\prime}\mathbf{w}_t + u_t \qquad u_t \sim \mathcal{N}\left(0, \sigma^2\right)$$

- Podemos reparametrizar el modelo definiendo un predictor $\mathbf{g}_t$ y un vector de parámetros $\boldsymbol{\theta}_t$, ambos de $p \times 1$ del siguiente modo:

$$\mathbf{g}_t = \begin{pmatrix} -1 \\ \mathbf{r}_t^e \end{pmatrix}, \quad \boldsymbol{\theta}_t = \begin{pmatrix} \alpha_t \\ \mathbf{w}_t \end{pmatrix}$$

para obtener

$$-r_{t,1} = \mathbf{g}_t'\boldsymbol{\theta}_t + u_t$$

- Este modelo es conocido en estadística bayesiana como *Dynamic linear model* y puede se resuelto introduciendo una dinámica para actualizar $\boldsymbol{\theta}_t$.

Proceso de aprendizaje

- De nuevo necesitamos una prior para $\boldsymbol{\theta}_t$. La prior se calcula con información disponible hasta $t-1$:
$$\boldsymbol{\theta}_t|I_{t-1} \sim \mathcal{N}\left(\mathbf{m}_{t|t-1}, \mathbf{V}_{t|t-1}\right)$$
- Aplicando el resultado sobre familias conjugadas, la posterior es
$$\boldsymbol{\theta}_t|I_t \sim \mathcal{N}\left(\mathbf{m}_t, \mathbf{V}_t\right)$$
con
$$\mathbf{m}_t = \mathbf{V}_t(-\sigma^{-2}\mathbf{g}_t r_{t,1} + \mathbf{V}_{t|t-1}^{-1}\mathbf{m}_{t|t-1})$$
y
$$\mathbf{V}_t = (\sigma^{-2}\mathbf{g}_t\mathbf{g}_t' + \mathbf{V}_{t|t-1}^{-1})^{-1}$$
- Un detalle: necesitamos un mecanismo de transición para pasar de $\boldsymbol{\theta}_t|I_t$ a $\boldsymbol{\theta}_{t+1}|I_t$. Usualmente se especifica una *ecuación de estado* que típicamente es un proceso Markoviano ($\boldsymbol{\theta}_{t+1} = \mathbf{A}\boldsymbol{\theta}_t + \mathbf{z}_t$). En el caso de portafolios grandes, la estimación de este proceso puede ser muy difícil.

Alternativa para transiciones

- En lugar de usar un filtro de Kalman, se puede introducir un factor de descuento para los datos del pasado, como un suavizamiento exponencial.
- Partiendo de una distribución inicial en $t = 0$:

$$\boldsymbol{\theta}_1 | I_0 \sim \mathcal{N}\left(\mathbf{m}_0, \mathbf{V}_0\right)$$

se puede aplicar un proceso iterativo para estimar los parámetros en T:

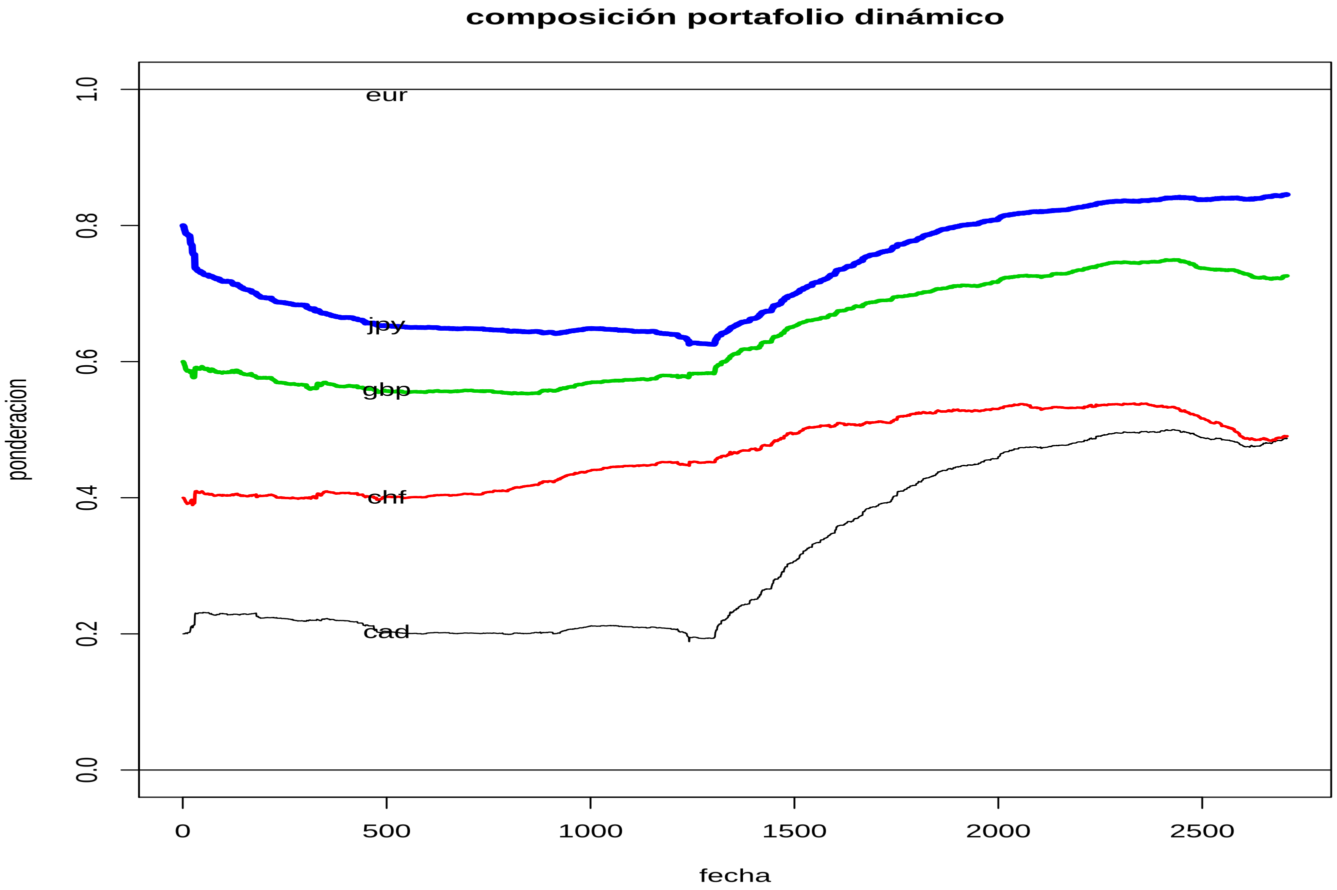
$$\mathbf{m}_T = \mathbf{V}_T(-\sigma^{-2}\sum_{t=1}^{T}(1-\delta)^{T-t}\mathbf{g}_t r_{t,1} + \mathbf{V}_0^{-1}\mathbf{m}_0)$$

y

$$\mathbf{V}_T = (\sigma^{-2}\sum_{t=1}^{T}(1-\delta)^{T-t}\mathbf{g}_t\mathbf{g}_t' + \mathbf{V}_0^{-1})^{-1}$$

Ejemplo

- El siguiente ejemplo fue corrido usando los parámetros $a = 0.001$, $\mathbf{m}_0$ y $\mathbf{V}_0$ como en el modelo bayesiano estático, suponiendo que el coeficiente α_t no está correlacionado con los pesos, y $\sigma = 0.002856$.



Falta mucho...

- Desarrollar un programa para el caso de varianza desconocida.
- Desarrollar criterios para proponer valores iniciales de a y σ^2, sin priores.
- Desarrollar un modelo que no permita posiciones cortas.
- Incorporar otras formas de modelar tracisiones en el modelo dinámico (sin el factor de descuento).
- $\dots$

Referencias

- Britten-Jones, M. 1997. The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights. Working Paper, London Business School.
- Britten-Jones, M. 1998. Statistical Portfolio Optimization Part 1: The Global Minimum Variance Portfolio Working Paper, London Business School.
- Ingersoll, Jonhatan E. Jr. 1987 *Theory of Financial Decision Making* Rowman and Littlefield.