

# Modelos Factoriales

Alfonso Novales  
Departamento de Economía Cuantitativa  
Universidad Complutense

Noviembre 2016  
Preliminary version  
No citar sin permiso del autor  
©Copyright 2014 A. Novales

## Contents

<b>1</b>	<b>Introducción</b>	<b>3</b>
<b>2</b>	<b>El modelo de un factor</b>	<b>4</b>
2.1	La diversificación de riesgo en el modelo de un factor . . . . .	6
2.2	Exponentially weighted moving average (EWMA) version of the one-factor model . . . . .	9
2.3	El modelo CAPM como modelo unifactorial . . . . .	12
<b>3</b>	<b>Evaluación de la bondad de un modelo factorial</b>	<b>12</b>
<b>4</b>	<b>Un modelo factorial general</b>	<b>13</b>
4.1	Ejemplos de factores de riesgo en mercados financieros . . . . .	15
4.2	Carteras . . . . .	15
4.3	Modelo de factores macroeconómicos . . . . .	16
4.4	Modelos de factores fundamentales . . . . .	18
4.4.1	El modelo Barra . . . . .	19
4.4.2	Factor-mimicking portfolio . . . . .	20
4.4.3	Los factores de Fama-French . . . . .	21
4.4.4	Modelo de factores ortogonales con varianza unitaria . . . . .	22
4.4.5	Rotación de factores . . . . .	24
<b>5</b>	<b>Modelos factoriales: Aplicaciones</b>	<b>25</b>
5.1	Style attribution analysis . . . . .	25
5.2	Multi-factor models in international portfolios . . . . .	26

<b>6</b>	<b>Identificación de tipos de interés a vencimientos específicos como factores de riesgo.</b>	<b>28</b>
6.1	A factor model of the term structure by regression . . . . .	29
6.2	Portfolio management . . . . .	32
6.3	A duration vector . . . . .	34
<b>7</b>	<b>Componentes principales</b>	<b>35</b>
7.1	El método de componentes principales . . . . .	36
7.1.1	Autovectores normalizados . . . . .	37
7.1.2	Elección del número de Componentes Principales . . . . .	39
7.1.3	Representación de los datos en función de Componentes Principales . . . . .	40
7.1.4	Risk decomposition: Systematic and idiosyncratic risks . . . . .	41
7.1.5	Componentes principales: una presentación alternativa . . . . .	42
7.1.6	Lack of scale-invariance in principal components (Mardia, Kent, Bibby) . . . . .	44
7.1.7	Hypothesis testing on eigenvalues and eigenvectors . . . . .	44
7.2	Algunas aplicaciones . . . . .	45
7.2.1	Estimation of fundamental factor models in a stock market . . . . .	45
7.3	Factor GARCH y GARCH ortogonal . . . . .	46
7.3.1	Otros ejemplos de aplicaciones de la técnica de componentes principales (omitir) . . . . .	47
<b>8</b>	<b>Applications of Permanent Components to Fixed Income management</b>	<b>48</b>
8.1	Componentes principales en el análisis de una estructura temporal de tipos de interés. . . . .	48
8.2	Principal components analysis of a set of interest rates . . . . .	50
8.3	Present value of a basis point: PV01 . . . . .	57
8.4	Approximations to PV01 . . . . .	58
8.5	Riesgo en carteras de renta fija . . . . .	59
8.6	Interest rate risk . . . . .	60
8.7	Summary of expressions . . . . .	62
8.8	Varias aplicaciones . . . . .	62
8.9	Open questions in the analysis of a term structure . . . . .	71
<b>9</b>	<b>Componentes principales en el análisis de mercados de valores</b>	<b>72</b>
9.1	Equity PCA factor models: [Case Study_II.2_PCA Equity factor Model] . . . . .	73
<b>10</b>	<b>Principal components and cointegration analysis</b>	<b>75</b>

# 1 Introducción

En muchas ocasiones, el inversor se enfrenta al análisis de la evolución temporal de un amplio conjunto de activos. Esto sucede, por ejemplo, cuando el inversor considera todas las empresas que cotizan en un mercado de valores con la intención de predecir o simular su comportamiento futuro sobre un determinado horizonte de inversión, pues dicho conjunto es generalmente demasiado amplio como para permitir su seguimiento estadístico. Lo mismo sucede con un mercado de renta fija, pues para anticipar los movimientos futuros en los precios de los bonos, el inversor debe tener un escenario para la evolución de las tasas de descuento a todos los plazos en los que los bonos que cotizan en el mercado pagan algún cupón.

Además, no cabe sino esperar que las rentabilidades de las acciones de distintas empresas, o los tipos de interés a distintos vencimientos, estén correlacionados, generando con ello un problema de colinealidad que, si es importante cuando se trabaja con un número reducido de variables, resulta aún mucho más problemática cuando el número de variables es elevado. La elevada correlación surge porque el número de fuentes diferentes de variabilidad que subyace al amplio conjunto de activos puede ser muy inferior al número de activos. La consecuencia es que el vector de rentabilidades de los activos tendrá una matriz de covarianzas aproximadamente singular.

Los métodos factoriales tratan de caracterizar precisamente un número reducido de fuentes de riesgo para un amplio conjunto de activos, ya sean de renta variable (en la forma de rentabilidades de acciones) o de renta fija, en la forma de tasas de descuento o de tipos de interés a distintos vencimientos. Los modelos de factores permiten representar la rentabilidad de los activos considerados en función de un número reducido de factores, teniendo para cada activo un vector de *betas*. Cada beta es la sensibilidad de la rentabilidad de un activo a cada uno de los factores de riesgo. Esto tiene mucha importancia. Por ejemplo, nos permite caracterizar carteras de renta variable con menor sensibilidad de un determinado factor de riesgo, ya sea el nivel de los tipos de interés, o el ciclo económico. También nos permite evitar repetir fuentes de riesgo en una determinada cartera, al no incluir en ella activos con unas betas similares. En un modelo de  $m$  factores para representar el riesgo en  $k$  activos tendremos una matriz  $k \times m$  de *betas*.

En otra línea de análisis, la elaboración de escenarios para un conjunto reducido de factores de riesgo mediante predicción o simulación puede ser un ejercicio muy manejable computacionalmente, mientras que sería prácticamente imposible realizarlo para los activos individuales. Además, una cuestión muy importante reside en hacer que dichos escenarios simulados futuros tengan las correlaciones observadas históricamente entre los activos considerados. Todo ello queda enormemente facilitado por los modelos de factores. No sólo son computacionalmente más simples, sino que permiten recuperar escenarios para el amplio conjunto de activos de los que partimos, manteniendo las propiedades estadísticas de dicho conjunto, a partir de escenarios para los factores.

En particular, la ganancia computacional se debe a que todos los cálculos que

precisan utilizar la matriz de varianzas y covarianzas de los activos considerados se simplifican enormemente mediante un modelo de factores. Estimar la matriz de varianzas y covarianzas de un número reducido de factores es muy sencillo, y de ella puede deducirse con sencillez la matriz de varianzas de los activos que componen la cartera.

Algunas aplicaciones de modelos factoriales:

- Identificar las principales fuentes de riesgo para una cartera de activos
- Cálculo del VaR: el valor en riesgo de una cartera requiere conocer la matriz de varianzas y covarianzas de las rentabilidades de los activos que la componen. Si dicho número de activos es elevado, como suele suceder, el cálculo puede ser muy complejo. La tarea queda muy simplificada utilizando la representación de los activos en función de los factores.
- Gestionar inversiones seleccionando carteras por el enfoque de Markowitz, dado que dicho enfoque precisa de la matriz de covarianzas de las rentabilidades de todos los activos disponibles, cuyo número será generalmente demasiado elevado.
- Construir carteras cubiertas contra las principales fuentes de riesgo
- Considerar escenarios para medir el riesgo de una cartera
- Facilitar los ejercicios de stress-testing, mucho más sencillos de realizar sobre los factores que sobre los activos individuales. ¿Qué es un shock o un escenario de extrema adversidad?

## 2 El modelo de un factor

El análisis de Markowitz para la selección de carteras eficientes se basa en la consideración de la rentabilidad esperada y la volatilidad de cada uno de los activos disponibles. Este análisis precisa conocer asimismo las covarianzas o, lo que es lo mismo, los coeficientes de correlación entre las rentabilidades de cada par de activos. Este requerimiento significa, en la mayoría de las situaciones de interés, que el analista necesita disponer de muchos parámetros estimados. Por ejemplo, en el caso de considerar la inversión en los 35 valores que configuran el Ibex35, el analista necesitaría: 35 rentabilidades esperadas, más 35 volatilidades, más  $35 \cdot 17$  correlaciones, un total de 665 parámetros.

Los modelos factoriales tienen como ventaja que reducen considerablemente la dimensionalidad del problema de gestión de carteras. El modelo de un factor postula:

$$r_{i_t} = a_i + \beta_i I_t$$

que la rentabilidad aleatoria de cada uno de los activos disponibles tiene dos componentes: uno,  $a_i$ , específico del activo, y otro,  $\beta_i I_t$ , que refleja un efecto común a todos los activos, debido al factor  $I_t$ , común a las rentabilidades de

todos ellos. Ambos factores son aleatorios, y las fluctuaciones que a lo largo del tiempo experimenta la rentabilidad de un activo se debe a variaciones en estos dos factores. El factor  $I_t$  podría ser la tasa de inflación,  $I_t = \pi_t$ , o la rentabilidad de un índice de mercado,  $I_t = r_{m_t}$ , por ejemplo.

Si extraemos la esperanza matemática de  $a_i$  tendremos:  $a_i = \alpha_i + u_i$ , donde  $\alpha_i$  es una constante específica de cada activo, mientras que  $u_i$  es el elemento aleatorio específico de cada uno de ellos. Por tanto, tenemos:

$$r_{i_t} = \alpha_i + \beta_i I_t + u_{i_t} \quad (1)$$

El modelo unifactorial está definido por dos características:

- a)  $Cov(u_{i_t}, I_t) = E(u_{i_t} \cdot I_t) = 0$ ,
- b)  $Cov(u_{i_t}, u_{j_t}) = 0$ .

La primera condición significa que el componente de la rentabilidad de un activo no explicado por el factor es independiente del valor que toma el factor. También significa que el grado en que el modelo unifactorial representa el comportamiento de la rentabilidad de un activo es independiente del valor numérico que toma el factor.

El residuo de dicha regresión se tomaría como el componente aleatorio, de media cero, específico de cada activo. Como es bien sabido dicho residuo está incorrelacionado con las variables explicativas de la regresión. Cabe observar dos cuestiones: en primer lugar, que un analista puede *construir empíricamente* tantos modelos de un factor como quiera. En segundo lugar, que un modelo de un factor no debe ser objeto de un análisis econométrico detallado, conducente a obtener sus *mejores estimaciones*.

La segunda hipótesis es la que realmente define el modelo, por cuanto que es, en general poco probable que se cumpla, al menos estrictamente. Hay que pensar que, considerando el caso de inversión en los 35 activos del Ibex35, tendríamos que comprobar que las 595 correlaciones que pueden obtenerse entre los residuos de las regresiones correspondientes a 2 activos distintos, fuesen todos ellos igual a cero. Por tanto, hay que tomar el modelo de 1 factor más como una construcción teórica que como una realidad empírica. Como modelo teórico, es muy útil para ayudarnos a pensar acerca del modo adecuado de medir el riesgo de un activo financiero, como veremos enseguida.

Esta segunda hipótesis significa que el factor es realmente informativo, pues recoge todo lo que de común hay en las fluctuaciones en las rentabilidades de los activos considerados. Dicho factor es la única razón que explica los movimientos conjuntos entre rentabilidades de activos. Pueden existir otros elementos además del factor, pero han de ser estrictamente específicos. Por ejemplo, supongamos que utilizamos como factor la rentabilidad de un índice de mercado. Podría pensarse que el elemento específico, es decir, el residuo que se obtiene en la regresión de la rentabilidad del activo sobre la rentabilidad de mercado, venga explicado por los dividendos distribuidos, y que estos tienen un comportamiento específico de cada activo. Pero puede ser difícil justificar tal interpretación: en este ejemplo, aun estando relativamente incorrelacionados

entre activos, los dividendos podrían estar muy correlacionados entre activos de un mismo sector (bancos, construcción, eléctricas, etc.).

A partir de ahora, por aclarar la notación, suponemos que se utiliza como factor la rentabilidad del índice de mercado. Si denotamos por  $Var(u_i) = \sigma_{u_i}^2$ ,  $i = 1, 2, \dots, N$ ,  $Var(I_t) = Var(r_m) = \sigma_m^2$ , bajo las hipótesis del modelo, tenemos:

$$\begin{aligned} r_{it} &= \alpha_i + \beta_i r_{mt} + u_{it} \\ r_{jt} &= \alpha_j + \beta_j r_{mt} + u_{jt} \\ E(r_i) &= \alpha_i + \beta_i E(r_m) \\ \bar{r}_i &= \alpha_i + \beta_i \bar{r}_m \\ \sigma_i^2 &= \beta_i^2 \sigma_m^2 + \sigma_{u_i}^2 \\ \sigma_{ij} &= \beta_i \beta_j \sigma_m^2 \end{aligned}$$

que muestran que: a) la rentabilidad esperada de un activo tiene dos componentes: uno debido a la rentabilidad esperada del mercado, y otro debido a  $\alpha_i$ , b) la varianza de la rentabilidad de un activo tiene asimismo dos componentes, uno nuevamente debido a la volatilidad de la rentabilidad del mercado, que se denomina componente sistemático, pues estará presente en todos los activos considerados, y otro debido a la volatilidad del componente específico  $u_{it}$ , de modo que el riesgo de un activo tiene un componente de riesgo *específico* o *idiosincrático* y otro componente de riesgo de mercado o *sistemático*, debido al factor común, c) por último, la covarianza entre la rentabilidad de dos activos se debe exclusivamente al riesgo de mercado.

En el modelo de un factor, el riesgo, medido como siempre por la desviación típica anualizada, puede expresarse:

$$\begin{aligned} Total\ variance &= Systematic\ variance + Specific\ variance \\ Total\ risk &= \sqrt{(Systematic\ risk^2 + Specific\ risk^2)} \end{aligned}$$

## 2.1 La diversificación de riesgo en el modelo de un factor

La representación anterior nos sirve para expresar el valor esperado y la varianza de la rentabilidad de una cartera formada por  $N$  activos, sin más que recordar que la expresión de su rentabilidad es:  $r_c = \sum_{i=1}^N X_i r_i$ , siendo  $X_i$  las ponderaciones de cada activo dentro de la cartera, por lo que tenemos:

$$\bar{r}_c = \sum_{i=1}^N X_i \bar{r}_i = \sum_{i=1}^N X_i \alpha_i + \sum_{i=1}^N X_i \beta_i \bar{r}_m = \alpha_c + \left( \sum_{i=1}^N X_i \beta_i \right) \bar{r}_m = \alpha_c + \beta_c \bar{r}_m$$

donde hemos *definido* el alfa y la beta de la cartera,  $\alpha_c$  y  $\beta_c$  a partir de los correspondientes a los activos que la componen, por:

$$\alpha_c = \sum_{i=1}^N X_i \alpha_i$$

$$\beta_c = \sum_{i=1}^N X_i \beta_i$$

Además,

$$\sigma_c^2 = \sum_{i=1}^N X_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=1, j < i}^N X_i X_j \sigma_{ij} = \sum_{i=1}^N X_i^2 \beta_i^2 \sigma_m^2 + 2 \sum_{i=1}^N \sum_{j=1, j < i}^N X_i X_j \beta_i \beta_j \sigma_m^2 + \sum_{i=1}^N X_i^2 \sigma_{u_i}^2$$

Estas expresiones muestran que podemos estimar la rentabilidad esperada y el riesgo de cualquier cartera a partir de estimaciones de<sup>1</sup>  $\alpha_i, \beta_i, \sigma_{u_i}^2, \bar{r}_m, \sigma_m^2$ , un total de  $3N + 2$  parámetros, frente a los  $2N + N(N - 1)/2$ . En el caso de los 35 valores del Ibex35, el número de parámetros requerido se reduce de 665 a 107, pero si seguimos un mercado amplio, la reducción es muy superior.

Es muy importante observar que el cálculo anterior se basa en el *supuesto de que los componentes específicos de las rentabilidades de los activos tienen correlación cero*. Cuando esto no sucede, la expresión anterior es tan solo una aproximación a la verdadera volatilidad de la cartera [ver Example II.1.2 de C. Alexander].

La volatilidad de una cartera puede también escribirse:

$$\begin{aligned} \sigma_c^2 &= \left( \sum_{i=1}^N \sum_{j=1}^N X_i X_j \beta_i \beta_j \right) \sigma_m^2 + \sum_{i=1}^N X_i^2 \sigma_{u_i}^2 = \left( \sum_{i=1}^N X_i \beta_i \right) \left( \sum_{i=1}^N X_j \beta_j \right) \sigma_m^2 + \sum_{i=1}^N X_i^2 \sigma_{u_i}^2 \\ &= \beta_c^2 \sigma_m^2 + \sum_{i=1}^N X_i^2 \sigma_{u_i}^2 \end{aligned}$$

Por tanto, si identificamos varianza con riesgo, tenemos que el riesgo de un activo individual es  $\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{u_i}^2$ , que depende en parte de su beta, y en parte de la varianza de su componente específico. Puesto que el efecto de  $\sigma_{u_i}^2$  sobre el riesgo de una cartera puede eliminarse al aumentar  $N$ , se conoce a dicho componente como *riesgo diversificable*, o *riesgo idiosincrático* del activo. Por el contrario,  $\beta_i^2 \sigma_m^2$  no disminuye al aumentar  $N$ , y  $\beta_i$  es la medida del componentes *no diversificable* del riesgo de un activo.

Consideremos ahora una cartera bien diversificada. Esta es una cartera que: *a)* invierte en un amplio conjunto de activos, *b)* no concentra el valor nominal invertido en un subconjunto reducido de los activos en los que invierte.

<sup>1</sup>También puede obtenerse a partir de  $\bar{r}_i, \sigma_{u_i}^2, \beta_i, \bar{r}_m, \sigma_m^2$ , nuevamente  $3N + 2$  parámetros, pues de los  $\bar{r}_i, i = 1, 2, \dots, n$  pueden obtenerse los  $\alpha_i, i = 1, 2, \dots, m$  sin ninguna dificultad.

En particular, una cartera de este tipo es una cartera que distribuye su valor nominal a partes iguales entre un conjunto de  $N$  activos, siendo  $N$  grande<sup>2</sup>. La volatilidad de dicha cartera es:

$$\sigma_c^2 = \beta_c^2 \sigma_m^2 + \sum_{i=1}^N \left( \frac{1}{N} \right)^2 \sigma_{u_i}^2 = \beta_c^2 \sigma_m^2 + \frac{1}{N} \left( \sum_{i=1}^N \frac{1}{N} \sigma_{u_i}^2 \right)$$

El término en paréntesis es el promedio de las varianzas de los términos específicos de las rentabilidades de los activos considerados. En términos empíricos, dicho término es el promedio de las varianzas residuales de las  $N$  regresiones estimadas en la construcción del modelo de un factor. Aunque su valor numérico variará con la muestra utilizada, no hay ninguna razón para que no sea estable.<sup>3</sup> Por tanto, su cociente por  $N$  tenderá a cero al aumentar el número de activos en la cartera. El componente del riesgo de una cartera que no puede eliminarse es el asociado al riesgo de mercado, por lo que nos queda:

$$\sigma_c \simeq \sqrt{\beta_c^2 \sigma_m^2} = \sigma_m \left( \sum_{i=1}^N X_i \beta_i \right)$$

Como  $\sigma_m$  es común a cualquier cartera que podamos construir, tenemos que *la contribución de un activo individual al riesgo de una cartera* se mide a través de su beta,  $\beta_i$ .

A *risk manager* uses the above expressions to construct historical time series for many different portfolios, so that she can compare their characteristics in relatively short histories, allowing for time variation in parameters (this is because the time series of historical portfolio returns will have been constructed maintaining constant the current weights). On the other hand, an *asset manager* will compare the characteristics of many individual assets using long histories (three to five years of monthly or weekly data).

Once we have estimated the one-factor model, we can proceed to compute *i*) historical mean returns as well as *ii*) their standard deviations and covariances, or their volatilities and correlations, and *iii*) the correlation matrix of the specific return components, and compare them with the value that would be obtained from the mean and variance market return and the estimated alphas and Betas, according to the expressions presented above.

**Example 1** *ExII.1.1: With weekly data from the beginning of January 2000 to end of August 2008 for NWL y Microsoft, las alfas estimadas son 0,358 (0,161), y -0,066(0,178) y las betas estimadas: 0,506 (0,071) y 1,104 (0,079). Los componentes de riesgo específico se estiman en 23,17% y 25,74%, respectivamente. Una cartera constituida en un 70% por NWL y un 30% por Microsoft, tendría un alfa de 0,231 y una beta de 0,685, con un riesgo específico de 23,97%. Esto*

<sup>2</sup> Puede apreciarse que la expresión anterior no se satisface exactamente para la cartera de mercado, lo que ilustra que las hipótesis del modelo de un índice son inconsistentes para el caso  $\sigma_c^2 = \sigma_m^2$ .

<sup>3</sup> Salvo que el componente específico de una rentabilidad no fuese estacionaria, lo cual generalmente no sucederá.



se obtiene utilizando las alfas y betas de los activos que componen la cartera, así como mediante una regresión de la rentabilidad de la cartera sobre la rentabilidad del índice.

En general, sin embargo, los componentes idiosincráticos del modelo de un factor no estarán incorrelacionados entre sí, por lo que los cálculos anteriores para estimar el riesgo idiosincrático pueden no ser correctos [como se muestra en el ejercicio II.1.2 del libro de C. Alexander]. En tal caso, el riesgo idiosincrático debería estimarse estimando una regresión con la rentabilidad de la cartera sobre la rentabilidad del índice, no agregando los riesgos idiosincráticos de los activos de la cartera.

**Example 2** *ExII.1.3: Suppose the total volatility of returns on a stock is 25%. A linear model with two risk factors indicates that the stock has betas of 0.8 and 1.2. The two factors have volatility of 15% and 20% and a correlation of -0.5. How much of the stock's volatility can be attributed to the risk factors, and how large is the stock's specific risk? R: Volatility due to the two factors is 20,78%, while specific risk is 13,89%.*

## 2.2 Exponentially weighted moving average (EWMA) version of the one-factor model

Risk management requires monitoring on a frequent basis (daily and even intraday) and parameter estimates must be left to vary to reflect current risk conditions. So we consider:

$$r_t = a_t + \beta_t I_t + u_t$$

The simplest possible way to estimate time varying parameters is through an Exponentially Weighted Moving Average mechanism (EWMA), using a smoothing constant  $\lambda$ :

$$\beta_{\lambda,t} = \frac{Cov_{\lambda}(r_t, I_t)}{Var_{\lambda}(I_t)}$$

$$Cov_{\lambda}(r_t, I_t) \equiv \sigma_{12_t} = (1 - \lambda)r_{t-1}I_{t-1} + \lambda\sigma_{12_{t-1}} = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} r_{t-i} I_{t-i}$$

$$Var_{\lambda}(I_t) \equiv \sigma_t^2 = (1 - \lambda)I_{t-1}^2 + \lambda\sigma_{t-1}^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} I_{t-i}^2$$

where we are assuming that the asset's return and the factor have zero expectation. A time varying correlation coefficient could similarly be defined by division of the covariance of both returns by the square root of the product of variances, both statistics defined as above. The value of  $\lambda$ , between 0 and 1, determines the persistence of the process of covariance or variance. A zero value

would produce immediate reactions to events, while a value close to one would make the variance or covariance almost constant. The higher the value of  $\lambda$ , the longer it will take for the effects on moments of events to die away. The EWMA mechanism is justified only if returns are *i., i.d.*

The value of  $\lambda$  can also be chosen to optimize a measure of fit, like the value of the log-likelihood function under Normality. It is sometimes chosen subjectively as it is the case with the 0.94 value used in Riskmetrics with daily data or the 0.97 value used with monthly data. A value of  $\lambda = 0.97$  amounts to a half-life of 23 days, close to one month. That is the length of time needed for the process to close half the initial distance to its long-run level.

*Exercise:* For assets of different nature, compute covariances and variances for alternative values of  $\lambda$ . Compare with moments computed with rolling windows of different length. Estimate the value of  $\lambda$ .

Under the EWMA specification, systematic risk is estimated by:

$$\text{Systematic Risk} = \sqrt{h} \hat{\beta}_{\lambda,t} \sqrt{\text{Var}_{\lambda}(I_t)}$$

where  $h$  denotes the number of returns per year, which will be around 250 when working with daily data. This analysis produces time varying betas and correlations. It is obviously interesting to observe the time changes in beta, one of the two components of systematic risk of the asset. Systematic risk will change over time as a function of changes in beta and changes in factor variance. Systematic risk may be low even for assets with beta above one if factor risk is high, and the opposite can also happen.

There is an interesting relationship between the equity beta and the relative volatility of the asset and the market:

$$\beta_{\lambda,t} = \rho_{\lambda,t} \sqrt{\frac{\text{Var}_{\lambda}(r_t)}{\text{Var}_{\lambda}(I_t)}} = \rho_{\lambda,t} \nu_{\lambda,t}$$

where  $\rho_{\lambda,t}$  is the correlation coefficient between the asset and the market and  $\nu_{\lambda,t}$  is relative volatility, both calculated with the weighted moments. A graphical comparison of time variation in correlation and relative volatility across assets will usually provide interesting information on the time evolution of beta-risk.

So,

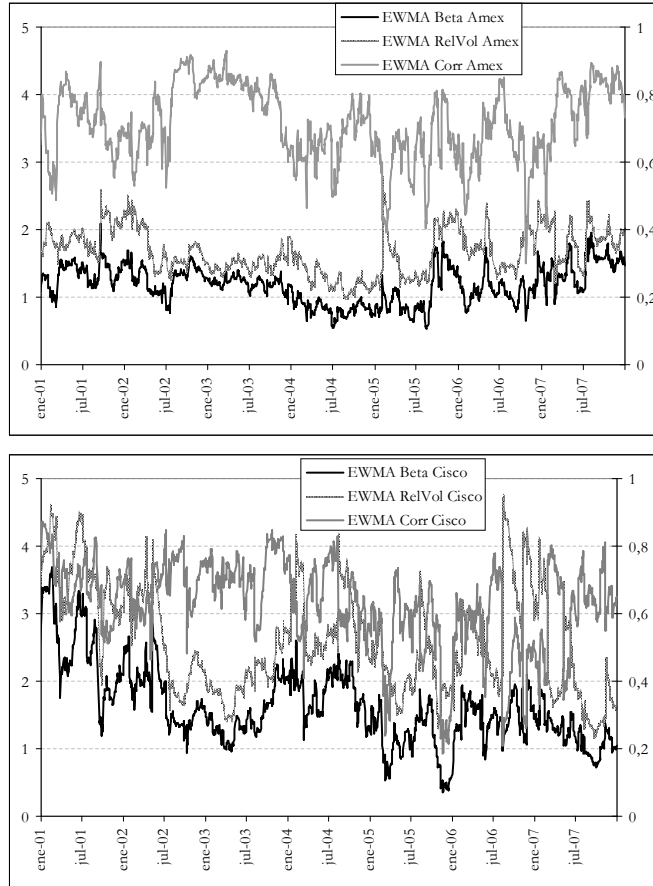
$$\text{Systematic Risk}_t \equiv SR_t = \sqrt{h} \rho_{\lambda,t} \nu_{\lambda,t} \sqrt{\text{Var}_{\lambda}(I_t)}$$

Si el riesgo sistematico aumenta por aumentos en la volatilidad del indice, podriamos cubrirnos con futuros sobre dicho indices, mientras que si se debe a aumentos en la volatilidad relativa, podriamos utilizar futuros sobre la volatilidad del activo, si los hubiera. Aumentos en la correlacion podrian sugerir utilizar futuros sobre el indice, asimismo.

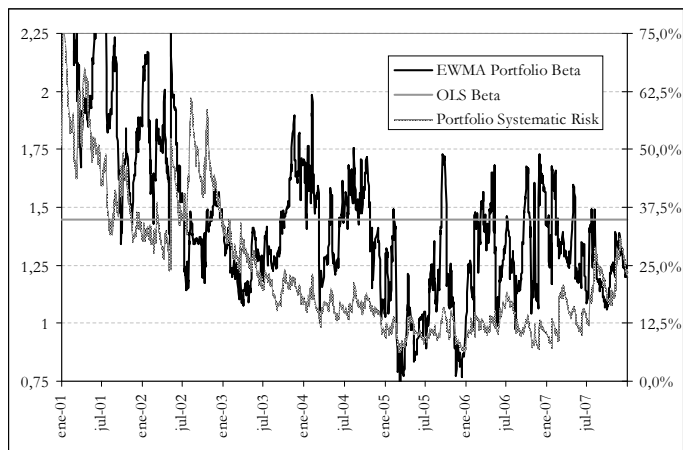
**Example 3** Example (Figures II.1.1 to II.1.3 in file Figures II.1). *The example considers a portfolio made up by Amex and Cisco. It is obvious that Cisco has a*

greater systematic risk than Amex (larger  $\beta$ ). The average market correlation is similar for Cisco and Amex, but Cisco is much more volatile than Amex, relative to the market and hence, EWMA correlation is much more unstable and Cisco beta is often considerably higher than Amex beta.

Beta, relative volatility and correlations for each asset



During 2001, the portfolio had a much higher  $\beta$  than the OLS  $\beta$  estimate, and the opposite happens in the second part of the sample. Systematic risk is the product of the portfolio  $\beta$  times the volatility of the market factor. This was low in the second part of the sample in spite of the fact that the  $\beta$  was larger than 1, because of the low volatility of the market factor. High systematic risk in the summer of 2002 was not due to a high  $\beta$  but to a high market volatility.



### 2.3 El modelo CAPM como modelo unifactorial

Un caso sin duda muy relevante de un modelo unifactorial es el modelo de mercado o Capital Asset Pricing Model (CAPM), que explica la rentabilidad de un activo, en exceso de la proporcionada por el activo sin riesgo, en función del diferencial que sobre esta última rentabilidad proporciona el índice de mercado,

#### Example 4

$$r_{it} - r_{Ft} = \alpha_i + \beta_i(r_{Mt} - r_{Ft}) + \varepsilon_{it}, \quad i = 1, 2, \dots, k; \quad t = 1, 2, \dots, T$$

*Tsay considera excesos de rentabilidad mensuales de 13 activos del S&P500, así como el propio índice, desde enero 1990 a diciembre 2003,  $k = 13, T = 168$ . Los activos son: Alcoa, E.G.Edwards, Caterpillar, Ford Motor, FedEx, General Motors, Hewlett-Packard, Kimberly-Clark, Mellon Financial, New York Times, Procter & Gamble, Chicago Tribune, Texas Instrument, y el propio índice. El archivo `m-fac9003.txt` contiene los excesos de rentabilidad mensuales de los activos en este orden. El tipo de interés de letras del Tesoro EEUU a 3 meses es utilizado como rentabilidad del activo sin riesgo. Los activos financieros y los de alta tecnología resultan tener betas y  $R^2$  más altos. El mayor  $R^2$  es de 0,41, indicando que la variabilidad específica es importante.*

## 3 Evaluación de la bondad de un modelo factorial

1. Las correlaciones entre factores comunes y específicos debe ser cero, aunque esto está garantizado si estimamos por mínimos cuadrados.
2. Las correlaciones entre componentes específicos deben ser cero; de lo contrario, sugerirían que hay un factor común adicional que no hemos recogido en el modelo.

3. Un modo de evaluar el ajuste de un modelo factorial es comparar las correlaciones entre rentabilidades que resultan del modelo, con las observadas en la muestra. Sin embargo, en el caso de un modelo unifactorial, dichas correlaciones serán todas igual a 1.
4. Otra alternativa es comparar la composición que surge de ambas matrices de correlaciones (la muestral y la del modelo de un índice) para una determinada cartera de referencia. Dicha cartera es la solución al problema:

$$\begin{aligned} \min_{\omega} \sigma^2(\omega) &= \omega' \Sigma \omega \\ \text{sujeto a } \omega' \mathbf{1}_k &= 1 \end{aligned}$$

donde  $\Sigma$  es la matriz  $k \times k$  de covarianzas entre las  $k$  rentabilidades, y  $\omega$  es un vector  $k \times 1$ .

La solución a este problema es:

$$\omega = \frac{\Sigma^{-1} \mathbf{1}_k}{\mathbf{1}'_k \Sigma^{-1} \mathbf{1}_k}$$

que habría que calcular con las rentabilidades originalmente observadas, así como con las rentabilidades ajustadas por el modelo factorial.<sup>4</sup> En el ejemplo considerado, las carteras de mínima varianza, calculadas de ambos modos, difieren notablemente en la ponderación asignada a uno de los valores.

5. Por último, es imprescindible evaluar el cumplimiento de las hipótesis que justifican el modelo. Concretamente, la ausencia de correlación entre los factores específicos. En el ejemplo mencionado de Tsay se observan algunas correlaciones por encima de 0,40.

## 4 Un modelo factorial general

Consideremos  $k$  activos y  $T$  observaciones temporales. Una representación general del modelo factorial es:

$$r_{it} = \alpha_i + \beta_{i1} f_{1t} + \beta_{i2} f_{2t} + \dots + \beta_{im} f_{mt} + \varepsilon_{it}, \quad t = 1, 2, \dots, T; \quad i = 1, 2, \dots, k, \quad \text{con } m \ll k$$

donde  $f_{jt}$ ,  $j = 1, 2, \dots, m$  representan los  $m$  factores comunes, y  $\beta_{ij}$  denotan las *cargas factoriales*, y  $\varepsilon_{it}$  es el componente específico del activo  $i$ , no explicado por los factores comunes.

Suponemos que el vector de factores  $f_t = (f_{1t}, f_{2t}, \dots, f_{mt})$  sigue un proceso estacionario  $m$ -dimensional con:

<sup>4</sup>Nótese que esta es la misma regla que vimos en el capítulo de Volatilidad. Cada ponderación se obtiene a partir de la suma de los elementos de una columna de  $\Sigma^{-1}$ , dividida por el agregado de dichas sumas.

$$\begin{aligned}
E(f_t) &= \mu_f \\
Cov(f_t) &= \Sigma_f, \text{ matriz simétrica } k \times k \text{ definida positiva} \\
E(\varepsilon_{it}) &= 0 \text{ para todo } i, t \\
Cov(f_{jt}, \varepsilon_{is}) &= 0 \text{ para todo } j, i, t, s \\
Cov(\varepsilon_{it}, \varepsilon_{js}) &= \sigma_i^2 \text{ si } i = j, t = s, \text{ e igual a } 0 \text{ en caso contrario}
\end{aligned}$$

por lo que las correlaciones entre activos son consecuencia de la presencia de los factores comunes  $f$ . Sin embargo, los factores comunes pueden estar correlacionados entre sí, con matriz de covarianzas  $\Sigma_f$ . Un supuesto habitual adicional es la ausencia de correlación serial en las variables  $r_t$  por lo que, de no ser así, el modelo debe aplicarse a los residuos de un modelo de series temporales que satisfagan las condiciones anteriores.

El modelo del activo  $i$ -ésimo puede escribirse:

$$r_{it} = \alpha_i + \beta'_i f_t + \varepsilon_{it}, i = 1, 2, \dots, k, t = 1, 2, \dots, T$$

donde  $\beta_i$  es el  $m$ -vector de los coeficientes del modelo anterior.<sup>5</sup> El modelo conjunto para los  $k$  activos es:

$$r_t = \alpha + B f_t + \varepsilon_t, t = 1, 2, \dots, T \quad (2)$$

con  $Cov(\varepsilon_t) = D = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2)$ , siendo  $B$  la matriz  $k \times m$  de coeficientes estimados.<sup>6</sup> La matriz de covarianzas del vector de rentabilidades es:

$$\text{Var}(r_t) = B \Sigma_f B' + D$$

donde  $\Sigma_f$  es la matriz  $m \times m$  de covarianzas de los factores. Nótese que estamos suponiendo que todos los momentos son invariantes en el tiempo. Cada elemento de la diagonal del primer sumando,  $B \Sigma_f B'$ , es la parte de la varianza de  $r_{it}$  explicada por los  $m$  factores comunes. Se conoce como la *comunalidad*, siendo específica de cada rentabilidad. El componente restante de la varianza del activo,  $\sigma_i^2$ , se conoce como varianza específica, o también como *uniqueness*.

<sup>5</sup>En la notación habitual de los modelos de regresión, tendríamos *para cada activo*  $i$ ,

$$R_i = \alpha_i 1_T + F \beta_i + E_i$$

donde  $R_i = (r_{i1}, r_{i2}, \dots, r_{iT})'$ ,  $F$  es una matriz  $T \times m$  cuya fila  $t$  es  $f'_t$ , y  $E_i = (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{iT})'$ , con matriz de covarianzas:  $\text{Var}(E_i) = \sigma_i^2 I_T$ .

<sup>6</sup>Si incluimos la constante como un factor más, el modelo (2) puede reescribirse:

$$r_t = \xi g_t + \varepsilon_t$$

donde  $g_t = (1, f_t)$ ,  $\xi = (\alpha, \beta)$ , una matriz  $k \times (m + 1)$ . Si trasponemos esta ecuación y apilamos todas las ecuaciones, tenemos:

$$R = G \xi' + E$$

donde  $R$  es una matriz  $T \times k$  cuya fila  $t$  es  $r'_t$  y cuya columna  $i$ -ésima es  $R_i$ ,  $G$  es una matriz  $T \times (m + 1)$  cuya fila  $t$  es  $g'_t$ , y  $E$  es una matriz  $T \times k$  de factores específicos cuya fila  $t$  es  $\varepsilon'_t$ .

Si los factores comunes  $f_t$  son observables, entonces la ecuación (2) es una regresión lineal multivariante. Este es el caso cuando utilizamos factores de riesgo macroeconómicos o financieros. Si reunimos el sistema de las  $k$  ecuaciones correspondientes a todos los periodos, tenemos:

$$r - \alpha \circ 1_T = (B \circ I_T) F + \varepsilon$$

donde  $r$  es  $kT \times 1$ ,  $F$  es  $mT \times 1$ ,  $B$  es  $k \times m$ ,  $B \circ I_T$  es  $kT \times mT$ ,  $\alpha$  es  $k \times 1$  y  $\varepsilon$  es un vector  $kT \times 1$ .

#### 4.1 Ejemplos de factores de riesgo en mercados financieros

- Sorpresas no esperadas en indicadores macroeconómicos, como el Índice de Precios de consumo, la oferta monetaria, el empleo, la tasa de paro. Se toman como sorpresas para cada variable los residuos obtenidos al estimar un modelo de predicción, ya sea un modelo ARIMA para cada variable, o un modelo VAR para todas ellas o para algún subconjunto de las mismas.
- Factores Fama-French para valoraciones de acciones. Ver página de Kenneth French
- Los tipos de interés a 3 meses y a 5 años como factores de riesgo para toda la estructura temporal de tipos cupón cero
- El nivel de los tipos de interés, la pendiente de la curva de tipos y su curvatura como tres factores de la estructura temporal.
- Las series temporales de parámetros estimados del modelo Nelson-Siegel como factores de riesgo para la estructura temporal de tipos cupón cero.

#### 4.2 Carteras

Recordemos que el modelo multifactorial es:

$$r_{it} = \alpha_i + \beta'_i f_t + \varepsilon_{it}, i = 1, 2, \dots, k, t = 1, 2, \dots, T$$

con componentes no explicados que tienen matriz de covarianzas:

$$\Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1k} \\ \dots & \dots & \dots & \dots \\ \sigma_{k1} & \sigma_{k2} & \dots & \sigma_k^2 \end{pmatrix}$$

In a completely specified factor model, this covariance matrix should be diagonal, as it has been discussed above. However, we are going to proceed at this point as if we are at an intermediate point of specifying the factor model, which is still incomplete, producing some correlations between the unexplained components of the vector of asset returns.

Consideremos una cartera constituida por posiciones en los  $k$  activos, con pesos:  $w' = (w_1, w_2, \dots, w_k)$ . The historical data on the portfolio, using 'current weights' is:

$$r_c = w'r$$

tendremos para cada período:

$$\begin{aligned} r_t &= \alpha + Bf_t + \varepsilon_t, \quad t = 1, 2, \dots, T \\ r_c &= w'r_t = w'\alpha + w'Bf_t + w'\varepsilon_t, \quad t = 1, 2, \dots, T \end{aligned}$$

por lo que el alfa de la cartera es  $\alpha_c = w'\alpha$ , su  $m$ -vector de betas es beta es  $\beta_c = w'B$ , and the portfolio's specific return is given by:  $\varepsilon_c = w'\varepsilon_t$ , the weighted sum of the assets' specific returns.

This expression for the portfolio's specific return shows the need to take into account the correlation between asset specific returns when estimating the specific risk of the portfolio. Assuming *lack of correlation* between factor returns and each asset's specific return, we have,

$$\text{Var}(r_c) = \beta'_c \Sigma_f \beta_c + w' \Omega w$$

displaying the three sources of risk: *i*) the risks coming from the portfolio's factor sensitivities  $\beta_c$ , *ii*) the risks of the factors themselves, represented by  $\Sigma_f$ , *iii*) the influence of the idiosyncratic risks of the assets in the portfolio, represented by  $w' \Omega w$ .

El componente de riesgo sistemático de la cartera (explicado por los factores) sería:  $\beta'_c \Sigma_f \beta_c$  y la volatilidad debida a los factores:  $\sqrt{\lambda \beta'_c \Sigma_f \beta_c}$  siendo  $\lambda$  el factor preciso para anualizar dicha volatilidad.<sup>7</sup> El riesgo idiosincrático sería igual a:  $\sqrt{\lambda(\text{Var}(r_c) - \beta'_c \Sigma_f \beta_c)}$ .

**Example 5** Example II.1.5: Consider a portfolio invested in three assets with weights: -0.25, 0.75 and 0.50, respectively. Each asset has a factor model representation with the same risk factors as in Example II.1.3, and betas: (0.2,1.2), (0.9, 0.2), and (1.3, 0.7). The two factors have volatilities: 15% and 20%, respectively, and their linear correlation coeficiente is -0,5. What is the volatility due to the risk factors (i.e., the systematic risk) for this portfolio? R: The portfolio factor betas are 1,275 and 0,20, and the volatility due to the two factors is 17,47%.

### 4.3 Modelo de factores macroeconómicos

En la modelización factorial de un mercado de renta variable, existen dos posibilidades: una es la búsqueda de combinaciones lineales de activos como factores, lo cual permite utilizar los datos de alta frecuencia (al menos diarios) de que habitualmente disponemos. Tal combinación lineal puede interpretarse como una cartera, cuyo valor de mercado es la combinación lineal de los precios de

<sup>7</sup>Que sería el número de observaciones muestrales contenidas en un período de un año.



los activos que la constituyen con los pesos dados por los coeficientes de la combinación lineal. Dichos coeficientes pueden normalizarse para que sumen la unidad. Hay que tener en cuenta, sin embargo, que salvo que se impongan restricciones en su cálculo, un factor de estas características tendrá pesos positivos y pesos negativos, por lo que para poder interpretarlo como una cartera debe ser posible que el potencial inversor pueda tomar posiciones cortas.

La segunda posibilidad, asimismo habitual, es la utilización de factores macroeconómicos, un índice de precios para incorporar el riesgo de inflación, y quizá el índice de producción industrial y el número de ocupados, el paro registrado o la tasa de paro como indicadores cíclicos. Todos estos indicadores son mensuales, lo que limita la frecuencia de los datos que podemos utilizar, y nos obliga a considerar el modo en que tomamos datos mensuales a partir de los datos financieros diarios, ya sea tomando el último dato de cada mes, el dato de mitad de mes, o la media de los datos diarios de dicho mes. Podríamos añadir además algunos indicadores financieros, como la rentabilidad de un índice de mercado, ya sea de renta fija o de renta variable, así como su volatilidad, el nivel de los tipos de interés, la pendiente de la curva de tipos. Esto se debe a la evidencia acerca de la sensibilidad de los mercados de renta variable a las expectativas acerca de las posibles fluctuaciones en el ciclo económico. Este enfoque fuerza a utilizar datos de menor frecuencia, debido a que los indicadores macroeconómicos no están disponibles en frecuencias por debajo de un mes. Pero precisamente esta propuesta pretende anticipar la evolución posible del mercado a medio plazo, por lo que renunciar al uso de datos diarios no es un gran problema. Una vez especificados los factores y habiendo estimado el modelo, la predicción acerca de la evolución del mercado de renta variable requerirá utilizar predicciones acerca de los valores futuros de los indicadores macroeconómicos utilizados. Alternativamente, puede hacerse una predicción por escenarios. En el caso de los indicadores macroeconómicos, esta posibilidad puede ser incluso más adecuada, ya que podemos tener una idea razonablemente aproximada acerca del comportamiento de la inflación, el crecimiento económico o el empleo durante el próximo año. La idea sería elaborar varios escenarios alternativos de evolución futura de los indicadores, obtener la predicción de la evolución del mercado de renta variable asociada (condicionada) a cada escenario, y promediar las distintas predicciones así obtenidas por la probabilidad que el analista asocie a cada escenario macroeconómico.

Si incluimos la constante como un factor más, el modelo (2) puede reescribirse:

$$r_t = \xi g_t + \varepsilon_t$$

donde  $g_t = (1, f_t)$  es un vector  $m \times 1$ , y  $\xi = (\alpha, \beta)$  es una matriz  $k \times (m + 1)$  de coeficientes. Si trasponemos esta ecuación y apilamos todas las ecuaciones, tenemos:

$$R = G\xi' + E$$

donde  $R$  es una matriz  $T \times k$  cuya fila  $t$  es  $r_t'$  y cuya columna  $i$ -ésima es  $R_i$ ,

$G$  es una matriz  $T \times (m+1)$  cuya fila  $t$  es  $g'_t$ , y  $E$  es una matriz  $T \times k$  de factores específicos cuya fila  $t$  es  $\varepsilon'_t$ .

En el modelo de factores macroeconómicos, los factores son observables. Habitualmente, tomaríamos un índice de precios para incorporar el riesgo de inflación, y quizá el índice de producción industrial y el número de ocupados, el paro registrado o la tasa de paro como indicadores cíclicos. Todos estos indicadores son mensuales, lo que limita la frecuencia de los datos que podemos utilizar, y nos obliga a considerar el modo en que tomamos datos mensuales a partir de los datos financieros diarios, ya sea tomando el último dato de cada mes, el dato de mitad de mes, o la media de los datos diarios de dicho mes. Podríamos añadir además algunos indicadores financieros, como la rentabilidad de un índice de mercado, ya sea de renta fija o de renta variable, así como su volatilidad, el nivel de los tipos de interés, la pendiente de la curva de tipos.

Podemos estimar el modelo por mínimos cuadrados:

$$\hat{\xi} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (G'G)^{-1}G'R$$

con residuos:

$$\hat{E} = R - G\hat{\xi}$$

y la matriz de covarianzas de  $\varepsilon_t$  se estima mediante:

$$\hat{D} = \text{diag} \left( \frac{\hat{E}'\hat{E}}{T - m - 1} \right)$$

donde  $\hat{D}$  se construye como matriz diagonal. El  $R^2$  del activo  $i$ -ésimo es:

$$R_i^2 = 1 - \frac{(\hat{E}'\hat{E})_{i,i}}{(R'R)_{i,i}}, \quad i = 1, 2, \dots, k$$

En este enfoque no hemos impuesto la ortogonalidad de los factores específicos entre sí. Por tanto, la estimación anterior no es eficiente, siendo teóricamente superior una estimación por mínimos cuadrados generalizados.

#### 4.4 Modelos de factores fundamentales

Este tipo de modelos utilizado fundamentalmente para mercados de renta variable, se basa en información sobre los "fundamentales" de las empresas, como la clasificación por industrias, el book-to market ratio, la capitalización, etc., o la clasificación por estilos: growth, value, para construir factores comunes que expliquen los excesos de rentabilidad. Hay dos versiones diferentes: el denominado enfoque BARRA, y el enfoque de Fama-French.

#### 4.4.1 El modelo Barra

En contraste con los enfoques anteriores, estimamos en cada periodo una regresión de sección cruzada para explicar las rentabilidades en exceso del conjunto de  $k$  activos, en diferencias respecto de sus respectivas medias muestrales, obteniendo un vector  $\hat{\beta}_t$  de sensibilidades:

$$\tilde{r}_t = \underset{k \times 1}{f'_t} \underset{k \times m}{\beta_t} + \underset{m \times 1}{\varepsilon_t}$$

Repetiendo para cada periodo, tendremos un conjunto de series temporales de betas y estas betas son en este enfoque los valores estimados de los factores. Nótese que algunas variables fundamentales serán constantes en el tiempo (la pertenencia de la empresa a un determinado sector) o variar solo ocasionalmente (el rating de la empresa); sin embargo, los factores  $\hat{\beta}_t$  irán variando al hacerlo las rentabilidades en exceso observadas.

En una primera etapa, estimamos para todos los periodos por Mínimos Cuadrados Ordinarios,

$$\hat{\beta}_{ot} = \underset{m \times 1}{(f_t \ f'_t)^{-1}} \underset{m \times 1}{(f \ \tilde{r}_t)}$$

y utilizamos los residuos para estimar:

$$\hat{D}_o = \text{diag} \left\{ \frac{1}{T-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}'_t \right\}$$

Dado que las rentabilidades estarán generalmente correlacionadas entre sí, procede estimar mediante Mínimos Cuadrados Ponderados (MCP):

$$\hat{\beta}_{gt} = \underset{m \times 1}{(f_t D^{-1} f'_t)^{-1}} \underset{m \times 1}{(f D^{-1} \tilde{r}_t)}$$

donde  $\tilde{r}_t$  denota la sección cruzada de rentabilidades en exceso en el instante  $t$  en desviaciones respecto de su media, y  $D = \text{cov}(\varepsilon_t)$  es una matriz  $k \times k$  que supondremos invariante en el tiempo.

Este estimador puede escribirse:

$$\hat{\beta}_{gt} = M \cdot \tilde{r}_t$$

con  $M = \left( f_t^{-1} \hat{D}_o^{-1} f'_t \right)^{-1} f_t^{-1} \hat{D}_o^{-1}$ , una matriz  $m \times k$ . Por tanto,  $\hat{\beta}_t$  será en cada periodo un vector  $m \times 1$  conteniendo los datos de los  $m$  factores para ese periodo. La matriz  $M$  describe el modo en que los factores fundamentales (recordemos que son las  $\hat{\beta}_t$ ), responden a cambios en las rentabilidades en exceso de cada activo.

Una vez obtenido el estimador, estimamos la matriz de covarianzas  $D$ , esta vez utilizando los residuos de la estimación MCP, obteniendo  $\hat{D}_g$ . Finalmente, la matriz de covarianzas de los factores es:

$$\hat{\Sigma}_{\hat{\beta}} = \frac{1}{T-1} \sum_{t=1}^T (\hat{\beta}_t - \bar{\beta}) (\hat{\beta}_t - \bar{\beta})'$$

donde:

$$\bar{\beta} = \frac{1}{T} \sum_{t=1}^T \hat{\beta}_t$$

y la matriz de covarianzas de las rentabilidades en exceso puede descomponerse:

$$Var(r_t) = \underset{k \times k}{f_t'} \hat{\Sigma}_{\hat{\beta}} \underset{k \times k}{f_t} + \underset{k \times k}{\hat{D}_g}$$

cuyo primer sumando puede considerarse cambiante en el tiempo o, lo que es más habitual, ser calculado utilizando los valores medios  $\bar{\beta}$ .

Un ejemplo es el denominado modelo de factores sectoriales, donde las variables  $f_{it}$  son variables dicotómicas que valen 1 si la empresa pertenece al sector  $i$ . La pertenencia a un sector no varía con el tiempo. La estimación MCO conduce a unos factores  $\hat{\beta}_{it}$  que son los promedios de las rentabilidades de las empresas de dicho sector.

#### 4.4.2 Factor-mimicking portfolio

En el contexto del modelo Barra, si consideramos un único "fundamental"  $f$ , entonces el factor fundamental tiene una clara interpretación. Tendríamos entonces el modelo factorial:

$$\underset{k \times 1}{\tilde{r}_t} = \underset{k \times 1}{f_t'} \underset{1 \times 1}{\beta_t} + \underset{k \times 1}{\varepsilon_t}$$

que conduce a una estimación de la varianza de las rentabilidades en exceso:

$$Var(r_t) = \hat{\sigma}_{\hat{\beta}}^2 \underset{k \times k}{(f_t \ f_t')} + \underset{k \times k}{\hat{D}}$$

La solución del problema de optimización:

$$\begin{aligned} & \min_w \left( \frac{1}{2} w' D w \right) \\ \text{suje}to \ a \quad & : \quad w' f = 1 \end{aligned}$$

es:

$$w' = (f D^{-1} f')^{-1} f D^{-1} \tag{3}$$

En efecto, formamos el Lagrangiano:

$$L(w, \lambda) = \frac{1}{2} w' D w - \lambda (w' f - 1)$$

con condiciones de optimalidad:

$$\begin{aligned} D w - \lambda f &= 0 \quad (k \text{ ecuaciones}) \\ w' f &= 1 \end{aligned}$$

Multiplicando por  $w'$  en

$$w' D w - \lambda w' f = 0 \Rightarrow \lambda = w' D w$$

y, por tanto, de la primera ecuación:

$$D w = (w' D w) f$$

y, sustituyendo, puede comprobarse fácilmente que tiene a (3) por solución:

$$\begin{aligned} D w &= D D^{-1} f' (f D^{-1} f')^{-1} = f' (f D^{-1} f')^{-1} \\ (w' D w) f &= (f D^{-1} f')^{-1} f D^{-1} D D^{-1} f' (f D^{-1} f')^{-1} f = (f D^{-1} f')^{-1} f \end{aligned}$$

Si normalizamos el vector  $w$  de modo que  $\sum_{i=1}^k w_i = 1$ , entonces la combinación lineal:

$$\hat{\beta}_t = w' r_t$$

define la cartera formada por los activos que están siendo considerados, que mejor reproduce el comportamiento del "fundamental"  $f$ . En efecto, la varianza del componente de dicha cartera no explicado por el "fundamental"  $f$  sería  $w' D w$ , y esto es precisamente lo que hemos minimizado con la elección de  $w'$  dada en (3). El vector  $w'$  definido en (3) es la matriz  $M$  en este caso particular.

De este modo puede caracterizarse la cartera que, en un determinado mercado, reproduce el factor inflación, o el factor tipos de interés, o el factor book-to-market.

#### 4.4.3 Los factores de Fama-French

Fama y French (1992) propusieron ordenar los activos de acuerdo con una característica fundamental como el book-to-market value, y considerar una cartera que toma posiciones largas en los activos del primer quintil (o tercio), y posiciones cortas en los activos del último quintil (o tercio). La rentabilidad observada en esta cartera es el factor fundamental, y el procedimiento se repite para cada variable. Finalmente, las betas se estiman mediante procedimientos de regresión. Fama y French detectaron la existencia de tres factores fundamentales que explican porcentajes significativos de las rentabilidades en exceso: a) el exceso

de rentabilidad del mercado, b) el resultado obtenido de carteras formadas por activos pequeños frente a las formadas por activos grandes (SMB, small minus big), c) el resultado obtenido de carteras formadas por activos de valor frente a activos de crecimiento (HML, high minus low).

Los activos "valor" son los que tienen un ratio alto entre su valor en libros y su valor de mercado. Otras veces la definición se basa en una alta rentabilidad por dividendo, o en un PER bajo. Los activos "growth" son los que ofrecen unas expectativas de crecimiento por encima del promedio, aunque para algunos (W. Buffet) no existen diferencias entre activos "valor" y activos "crecimiento".

#### 4.4.4 Modelo de factores ortogonales con varianza unitaria

En ocasiones, se busca un modelo de factores ortogonales, es decir un vector  $f_t$  con  $E(f_t) = 0_m$ , y  $\Sigma_f = I_m$ . Este es el caso del modelo de componentes principales que analizaremos más adelante. Que la esperanza matemática de los factores sea cero no es restrictivo pues podemos restar a cada variable original su media muestral, como estimador consistente de la esperanza matemática de la variable, para agregársela posteriormente. Que las varianzas sean unitarias se consigue aplicando el factor de escala apropiado a la serie temporal de cada factor o tipificándolas.<sup>8</sup> Trabajando con componentes principales, éste será habitualmente el caso.

Supongamos que, en efecto, los factores tienen varianza unitaria ( $\Sigma_f = I_m$ ). Como en todo modelo factorial la matriz de covarianzas  $Cov(\varepsilon_t) = D$  ha de tener (teóricamente) estructura diagonal, tendremos:

$$\begin{aligned} Var(r_t) &= BB' + D \text{ matriz } k \times k \\ Cov(r_t, f_t) &= E(r_t f_t) = BE(f_t f_t') + E(\varepsilon_t f_t) = B, \text{ matriz } k \times m \end{aligned}$$

Por tanto, en el modelo de factores ortogonales, tenemos:

$$\begin{aligned} Var(r_{it}) &= \beta_{i1}^2 + \dots + \beta_{im}^2 + \sigma_i^2, \quad i = 1, 2, \dots, k \\ Cov(r_{it}, r_{jt}) &= \beta_{i1}\beta_{j1} + \dots + \beta_{im}\beta_{jm}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, k \\ Cov(r_{it}, f_{jt}) &= \beta_{ij}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, m \\ Corr(r_{it}, f_{jt}) &= \beta_{ij} / \sqrt{\beta_{i1}^2 + \dots + \beta_{im}^2 + \sigma_i^2}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, m \end{aligned}$$

En este caso, la cantidad  $c_i^2 = \beta_{i1}^2 + \dots + \beta_{im}^2$ , la suma de los cuadrados de las cargas (*loadings*) de la variable  $i$  sobre los  $m$  factores comunes, es la parte de la varianza de  $r_{it}$  explicada por los  $m$  factores comunes. Se conoce como la *comunalidad*, siendo específica de cada rentabilidad. El componente restante de la varianza del activo,  $\sigma_i^2$ , se conoce como varianza específica, o también como *uniqueness*. La varianza de la variable  $i$  puede escribirse:

<sup>8</sup>Es decir, restando de cada dato su media muestral y dividiendo la diferencia por la desviación típica muestral.

$$\text{Var}(r_{it}) = c_i^2 + \sigma_i^2$$

No toda matriz de covarianzas admite una representación con factores ortogonales.

**Estimación del modelo estadístico factorial** Recordemos que en este enfoque, ni los factores ni las cargas son observables. El modelo factorial estadístico puede estimar bien por el procedimiento de Componentes Principales, o bien por Máxima Verosimilitud. El primero tiene la ventaja de no precisar ningún supuesto acerca de la distribución de probabilidad de los datos, si bien la solución es siempre aproximada. Para ello, calculamos los  $m$  primeros CPs,  $e_j, j = 1, \dots, m$ , de la matriz de covarianzas/correlaciones muestrales  $\hat{\Sigma}_r$ , que utilizamos como factores comunes, y tomamos como cargas factoriales:

$$\hat{\beta} = [\hat{\beta}_{ij}] = [\sqrt{\hat{\lambda}_1} \hat{e}_1 \quad \sqrt{\hat{\lambda}_2} \hat{e}_2, \dots, \sqrt{\hat{\lambda}_m} \hat{e}_m]$$

Una vez estimado el modelo, estimamos la varianza específica de cada variable mediante:

$$\hat{\sigma}_i^2 = \hat{\sigma}_{ii,r} - \sum_{j=1}^m \hat{\beta}_{ij}^2$$

donde  $\hat{\sigma}_{ii,r}$  denota el elemento  $i$ -ésimo en la diagonal de  $\hat{\Sigma}_r$ , y estimamos su comunalidad mediante:

$$\hat{c}_i^2 = \hat{\beta}_{i1}^2 + \hat{\beta}_{i2}^2 + \dots + \hat{\beta}_{im}^2$$

Si formamos la matriz diagonal  $D$  con los elementos  $\hat{\sigma}_i^2 : D = (\hat{\sigma}_i^2)$ , el error de aproximación puede evaluarse mediante la diferencia:

$$\hat{\Sigma}_r - (\hat{\beta} \hat{\beta}' + \hat{D})$$

que tendrá ceros en su diagonal principal, pero no fuera de ella. La suma de los cuadrados de los elementos de la matriz diferencia está acotada superiormente por  $\hat{\lambda}_{m+1}^2 + \dots + \hat{\lambda}_k^2$ . Un aspecto adicional de este procedimiento es que, por construcción, las cargas en los factores no varían si añadimos factores adicionales.

Para estimar por MV necesitamos un supuesto acerca de la distribución de probabilidad conjunta de los factores comunes  $f_t$  y específicos,  $\varepsilon_t$ . Si por ejemplo, los factores son Normales, entonces  $r_t$  seguirá una distribución Normal multivariante con esperanza  $\mu$  y matriz de covarianzas  $\Sigma_r = \beta \beta' + D$ . La estimación es computacionalmente delicada, por lo que puede fijarse  $\mu$  igual al vector de medias muestrales. También puede imponerse la restricción de que el producto  $\beta' D^{-1} \beta$  es una matriz diagonal. Al estimar por MV hay que imponer un determinado número de factores comunes, por lo que el número de ellos puede contrastarse mediante un test de Razón de Verosimilitudes:

$$LR(m) = \left( T - 1 - \frac{1}{6}(2k + 5) - \frac{2}{3}m \right) \left( \ln |\hat{\Sigma}_r| - \ln |\hat{\beta}\hat{\beta}' + \hat{D}| \right)$$

que, bajo la hipótesis nula de que existen  $m$  factores comunes, se distribuye asintóticamente como una  $\chi^2$  con  $\frac{1}{2}((k - m)^2 - k - m)$  grados de libertad.

#### 4.4.5 Rotación de factores

Por otra parte, la representación mediante factores ortogonales, cuando existe, no es única, pues si tomamos una matriz  $P$  ortogonal ( $PP' = P'P = I_m$ ) y construimos  $B^* = BP, F^* = P'F$ , tenemos:

$$r - \alpha \circ 1_T = (B \circ I_T)F + \varepsilon = (BP \circ I_T)P'F + \varepsilon = B^*F^* + \varepsilon$$

Además:  $E(F^*) = 0, Var(F^*) = P'Var(F)P = P'P = I_m$ . Por tanto,  $B^*$  y  $F^*$  configuran otra representación en factores ortogonales. La ausencia de unicidad de los factores es a la vez una debilidad y una ventaja del procedimiento. Como ventaja, nos permite llevar a cabo *rotaciones* que pueden facilitar la interpretación de los factores.

En particular, si consideramos una matriz ortogonal  $P$ , de orden  $m \times m$ , tenemos:

$$r_t - \mu = \beta f_t + \varepsilon_t = \beta PP' f_t + \varepsilon_t = \beta^* f_t^* + \varepsilon_t$$

y es fácil ver que si antes se cumplían las condiciones expuestas, la nueva descomposición, con  $\beta^* = \beta P$  y  $f^* = P'f$  también las cumple. Esto quiere decir que las cargas factoriales estimadas (*loadings*) son relativamente arbitrarias, y tenemos la posibilidad de buscar una rotación de los factores,  $f_t^* = P'f_t$ , que sea interpretable.

Las communalidades y las varianzas específicas son invariantes a la rotación de factores:

$$\beta^* \beta^{*'} + D = \beta PP' \beta' + D = \beta \beta' + D$$

Un criterio popular de rotación de factores es el conocido como *varimax*, consistente en seleccionar la matriz ortogonal  $P$  de modo que maximice la cantidad:

$$Max_P V = \frac{1}{k} \sum_{j=1}^m \left[ \sum_{i=1}^k \tilde{\beta}_{ij}^{*4} - \frac{1}{k} \left( \sum_{i=1}^k \tilde{\beta}_{ij}^{*2} \right)^2 \right]$$

donde  $\tilde{\beta}_{ij}^* = \beta_{ij}^*/c_i$ , siendo  $c_i$  la commonality del activo  $i$ -ésimo. Como puede verse, el corchete viene a ser  $k$  veces la varianza de las  $\tilde{\beta}$  al cuadrado, por lo que esta rotación persigue maximizar dicha varianza. Por tanto, se trata de encontrar grupos de coeficientes grandes y otros despreciables en cada columna de la matriz de cargas factoriales rotada, lo cual significaría que hemos rotado los factores de manera que solo unos pocos aparecen como importantes para explicar cada variable  $r_i$ .



## 5 Modelos factoriales: Aplicaciones

### 5.1 Style attribution analysis

Using some specific assets (indices or portfolios) as references for factors, it is interesting to estimate the management style of a given portfolio. This is important to evaluate fund management, which is sometimes subject to some regulations or to some indications from the financial director of the management institution. Specially important is this analysis to evaluate the management and performance of *hedge funds*.

Factors for style management may be stock market indices, bond indices, or interbank or many market rates. But they can also be indexes of *value* or *growth* stocks. To evaluate hedge funds, additional factors may include option price indices, exchange rates, credit spreads. A *value* stock is one that trades below book value. The asset value per share is high relative to the share price, and we could expect the price to raise. Usually its price-earnings ratio will be below the market average. A *growth* stock is one with a lower than average price-earnings to growth ratio. The rate of growth of earnings is high relative to its price-earnings ratio. These appear attractive because of potential growth in the firm assets. Value indicators may include the *book-to-price ratio* and the *dividend yield*, while growth indicators may include the *growth in earnings per share* and the *return on equity*. Value and growth style indices exist from different financial institutions: S&P 500 value index, S&P 500 growth index, Russell 1000 value index, etc.. As the number of firms increase, their average market capitalization decreases. So the S&P500 value index contains value stocks with an average market capitalization much higher than those firms in the Wilshire 5000 value index.

**Remark 6** References: *These indices can be downloaded from Yahoo! Finance. Look at the left side menus of: <http://finance.yahoo.com/funds> (Funds by family, Top performers)*

Obviously, the choice of style factors is very important for the results. We should include a large enough family representing the basic asset classes which are relevant to the portfolio being analyzed, without adding up too much collinearity. Style attribution is based on solving the problem:

$$\begin{aligned} & \min_{\beta} (y - X\beta)^2 \\ \text{subject to} & : \sum_{i=1}^k \beta_i = 1, \beta_i \geq 0, i = 1, 2, \dots, k \end{aligned}$$

**Example 7** Example II.1.4: *Perform style attribution on the mutual funds: VIT: Vanguard Index Trust 500 index, FAA: Fidelity Advisor Aggressive Fund, FID: Fidelity Main Mutual Fund, using as style factors: a) Russell 1000 value: mid cap, value factor, b) Russell 1000 growth: mid cap, growth factor, c) Russell*

2000 value: small cap, value factor, d) Russell 2000 growth: small cap, growth factor.

## 5.2 Multi-factor models in international portfolios

Consider an investment in a single foreign asset. The exchange rate is defined as the number of units of the domestic currency that must be delivered for each unit of the foreign currency. The log returns satisfy:

$$P_t^d = P_t^f E \Rightarrow R_d = R_f + X = \beta R + X$$

where we have assumed a single factor representation for foreign market risk. Hence, there are two risk factors affecting the domestic return on the asset: a) the exchange rate, with a beta of 1, and b) the foreign market index or risk factor, with a beta of  $\beta$ . Therefore,

$$\text{Systematic variance} = \text{Var}(\beta R + X) = \beta^2 \text{Var}(R) + \text{Var}(X) + 2\beta \text{Cov}(R, X)$$

and the systematic variance has three components: a) the equity variance,  $\beta^2 \text{Var}(R)$ , b) the forex variance,  $\text{Var}(X)$ , c) the equity-forex covariance:  $2\beta \text{Cov}(R, X)$ . If we have a portfolio of assets in the same class, we will have a similar decomposition, with the beta of the portfolio being related to the betas of individual assets in the standard fashion.

To generalize the analysis, let us consider a large international portfolio with exposure to  $k$  different countries. For simplicity, we assume that there is a single market risk factor in each country. We denote by  $R_1, \dots, R_k$  the returns to the market risk factors, by  $\beta_1, \dots, \beta_k$  the portfolio betas with respect to each market factor, and by  $X_1, \dots, X_k$  the foreign exchange rates. Assuming  $R_1$  is the domestic risk factor, then  $X_1 = 1$  and there are  $k$  equity risk factors, but only  $k-1$  foreign exchange risk factors. Let  $w = (w_1, \dots, w_k)'$  be the country portfolio weights. The systematic return on the portfolio is:

$$w_1 \beta_1 R_1 + w_2 (\beta_2 R_2 + X_2) + \dots w_k (\beta_k R_k + X_k) = (Bw)'x \quad (4)$$

where  $x$  is the  $(2k-1) \times 1$  vector of equity and forex risk returns, and  $B$  is the  $(2k-1) \times k$  matrix of risk factor betas:

$$x = (R_1, R_2, \dots, R_k, X_2, \dots, X_k)' \quad \text{and} \quad B = \begin{pmatrix} \beta_1 & 0 & 0 & \dots & 0 \\ 0 & \beta_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \beta_k \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Taking variances in (4) we get:

$$\text{Systematic variance} = (Bw)' \Omega (Bw)$$

where  $Bw$  is a  $(2k - 1) \times 1$  vector:  $Bw = [\beta_1 w_1, \beta_2 w_2, \dots, \beta_k w_k, w_2, \dots, w_k]$ , and  $\Omega$  is a  $(2k - 1) \times (2k - 1)$  matrix:

$$\Omega = \begin{pmatrix} \Omega_E & \Omega_{EX} \\ \Omega'_{EX} & \Omega_X \end{pmatrix}$$

where  $\Omega_E$  is the  $k \times k$  covariance matrix of the equity risk factor returns,  $\Omega_X$  is the  $(k - 1) \times (k - 1)$  covariance matrix of the forex risk returns, and  $\Omega_{EX}$  is the  $k \times (k - 1)$  'quanto' covariance matrix containing the cross-covariances between the equity risk factor returns and the forex risk factor returns:

$$\Omega = \begin{pmatrix} \text{Var}(R_1) & \text{Cov}(R_1, R_2) & \dots & \text{Cov}(R_1, R_k) & \text{Cov}(R_1, X_2) & \dots & \text{Cov}(R_1, X_k) \\ \text{Cov}(R_1, R_2) & \text{Var}(R_2) & & \text{Cov}(R_2, R_k) & \text{Cov}(R_2, X_2) & \dots & \text{Cov}(R_2, X_k) \\ & & \dots & & & & \\ \text{Cov}(R_1, R_k) & & & \text{Var}(R_k) & \text{Cov}(R_k, X_2) & & \text{Cov}(R_k, X_k) \\ \text{Cov}(R_1, X_2) & & & \text{Cov}(R_k, X_2) & \text{Var}(X_2) & \dots & \text{Cov}(X_2, X_k) \\ & & & & & \dots & \\ \text{Cov}(R_1, X_k) & & & \text{Cov}(R_k, X_k) & \text{Cov}(X_k, X_2) & & \text{Var}(X_k) \end{pmatrix}$$

The 'quanto' covariance matrix can be either positive or negative. In the latter case, the systematic variance will be less than the sum of the equity variance and the forex variance.

Particionando  $Bw = (\tilde{\beta}, \tilde{w})$ , con  $\tilde{w} = (w_2, \dots, w_k)$ ,  $\tilde{\beta} = w' \text{diag}(\beta_1, \dots, \beta_k) = (w_1 \beta_1, \dots, w_k \beta_k)'$  tenemos:

$$(Bw)' \Omega = \left( \tilde{\beta}' \Omega_E + \tilde{w}' \Omega_{EX}, \tilde{\beta}' \Omega_{EX} + \tilde{w}' \Omega_X \right) \Rightarrow (Bw)' \Omega (Bw) = \left( \tilde{\beta}' \Omega_E \tilde{\beta} + 2 \tilde{\beta}' \Omega_{EX} \tilde{w} + \tilde{w}' \Omega_{EX} \tilde{\beta} + \tilde{w}' \Omega_X \tilde{w} \right)$$

Hence, we can decompose systematic variance into its *equity*, *forex* and *equity-forex* components, as:

$$\text{Systematic variance} = w' (B' \Omega B) w = \tilde{\beta}' \Omega_E \tilde{\beta} + \tilde{w}' \Omega_X \tilde{w} + 2 \tilde{\beta}' \Omega_{EX} \tilde{w}$$

**Example 8** Example II.1.6: A UK investor holds 2.5 million pounds in UK stocks with a FTSE100 market beta of 1.5, 1 million pounds in US stocks with an S&P500 market beta of 1.2, and 1.5 million pounds in German stocks with a DAX30 market beta of 0.8. The volatilities and correlations of the FTSE100, S&P500 and DAX30 indices, and the USD/GBP and EUR/GBP exchange rates are estimated. Calculate the systematic risk of the portfolio and decompose it into equity and forex and equity-forex components.  $R$ :

## 6 Identificación de tipos de interés a vencimientos específicos como factores de riesgo.

The term structure of interest rates (TSIR) is a curve made up by zero-coupon curve rates for a large number of maturities. In some markets, like secondary debt markets, we need to estimate an analytical model that provides us with a continuous representation of zero coupon interest rates across the maturity range considered. In some other cases, like interbank markets, or interest rate swaps, we have already zero coupon rates observed at fixed maturities. There may be a fairly large number of maturities, as in swap markets, or just a few of them, as in markets for interbank deposits. Trading in such markets, or managing a fund in them requires evaluating the risk associated to each maturity, but there may be a large number of them. It makes sense then to try to summarize the time fluctuations in a large of maturities by those in a short number of them. This is a crucial aspect of risk management.

There are alternative approaches that can be followed to characterize risk factors across a term structure of interest rates. We could search for risk factors in the form of macroeconomic variables: inflation, growth, oil prices, and so on. Unfortunately, these models do not show a clear superiority over simpler alternatives in empirical tests is usually explained by the somewhat arbitrary choice of proxies for the risk factors. That a reduced number of factors can capture the uncertainty in changes in the TSIR is an interesting issue, since many equilibrium fixed income valuation models assume that bond prices are a function of a small number of *state variables* that follow a diffusion process [Vasicek, O.A. (1977), "An equilibrium characterization of the term structure", Journal of Financial Economics, 5, 177-188, Cox, J., Ingersoll, J., and S. Ross, (1985), "A theory of the term structure of interest rates", Econometrica, 53, 385-408] In these models, interest rates at pre-specified maturities are used as proxies for the unobserved state variables.

Estos trabajos sugieren la posibilidad de utilizar el análisis de regresión para tratar de caracterizar factores de riesgo, entendidos como tipos de interés a vencimientos determinados. Este es el enfoque que seguimos en esta sección. Para ello, habría que partir de un vector bastante amplio de tipos de interés, posiblemente deducidos de la estimación de curvas cupón cero según uno de los distintos modelos propuestos en la literatura: *splines* de Vasicek, McCulloch, Nelson-Siegel, Svensson, etc.. En este análisis no es difícil encontrar dos vencimientos, o quizá tres, que explican un elevado porcentaje de las fluctuaciones en los tipos de interés a todos los vencimientos. Para ello, se estiman regresiones para explicar las *variaciones* en los tipos de interés (para tener variables estacionarias) a cada vencimiento por medio de posibles subconjuntos de factores, que serían las variaciones en uno, dos o tres tipos de interés. No se parte de ningún *a priori* acerca de los vencimientos que contienen más información, por lo que este análisis requiere la estimación de un alto número de regresiones. Puede probarse que la suma de los coeficientes de determinación obtenidos con cada conjunto de factores posible es un buen criterio para discernir entre los fac-

tores con mayor capacidad explicativa. Pero conviene ponderar los coeficientes de determinación por la varianza de la variación en tipos que se trata de explicar en cada regresión.

Así habremos encontrado el factor que, utilizado por sí solo, explica mejor el comportamiento de todo el conjunto de tipos de interés seleccionados. Si queremos encontrar no uno, sino dos factores, estimamos conjuntos de regresiones como antes, pero esta vez con un par constante de tipos de interés como variables explicativas. Se prueba con cada par posible de tipos de interés como variables explicativas, y se toman como factores de riesgo el par de tipos que produjo la mayor suma de  $R^2$  corregidos. Tendríamos así también una estimación de las sensibilidades de cada tipo a cada factor de riesgo. Representadas gráficamente, veremos que los tipos adyacentes a cada factor tienen una alta sensibilidad a éste, y mucho menor a los restantes.

Una vez caracterizados los factores de riesgo (que serán tipos de interés a determinados vencimientos) si ajustamos una curva<sup>9</sup> a las sensibilidades estimadas para los vencimientos considerados, con respecto a un determinado factor, podríamos derivar de dicha regresión sensibilidades estimadas respecto a dicho factor para los tipos de intereses a cualquier vencimiento, incluso si no se observan en el mercado. Esto es importante en la valoración de activos de renta fija, para lo que cada flujo derivado del mismo se descuenta de acuerdo con el tipo de interés correspondiente al período desde del día de la valoración, hasta la percepción del flujo.

Así, pueden diseñarse escenarios de variaciones en precios de activos de renta fija. Lo importante es que, para ello, basta con considerar un determinado escenario únicamente para los factores de riesgo. Dicho escenario implica, a través de las sensibilidades estimadas, un escenario para cada uno de los tipos utilizados en la valoración del activo y, con ello, un escenario de precio del mismo. Si se elaboran varios escenarios, asignando una determinada probabilidad a cada uno de ellos, puede procederse de modo similar. Este es un problema de gestión de carteras de renta fija.

## 6.1 A factor model of the term structure by regression

In this approach, spot rates at different maturities are used as potential risk factors, breaking the arbitrariness mentioned above in the implementation of equilibrium valuation models. Elton, E.J., M.J. Gruber and R. Michaely [(1990) "The structure of interest rates and immunization", *Journal of Finance* 45, 629-642] assume that unexpected changes in interest rates are linearly related to two unknown factors  $F_1$  and  $F_2$ ,

$$dr_{i,t} = \beta_{i,0} + \beta_{i,1}dF_{1,t} + \beta_{i,2}dF_{2,t} + u_{i,t}, t = 1, 2, \dots, T, i = 1, 2, \dots, n \quad (5)$$

---

<sup>9</sup>Por ejemplo, un polinomio en los vencimientos correspondientes a los tipos de mercado observados.

The starting point is a large set of interest rates,  $\{r_{i,t}\}_{i=1,\dots,n}$ . In this approach, the factors will be identified with specific interest rates. Since interest rate mean changes are close to zero, the model is specified as,

$$dr_{i,t} = a_i dr_{z,t} + b_i (dr_{x,t} - dr_{z,t}) + u_{i,t} \quad (6)$$

where it is clear that any other number of factors could also be considered. It can be alternatively written as,

$$dr_{i,t} = (a_i - b_i) dr_{z,t} + b_i dr_{x,t} + u_{i,t} \quad (7)$$

The coefficient of determination when explaining  $dr_{i,t}$  by  $dr_{z,t}$  and  $dr_{x,t} - dr_{z,t}$  is given by,

$$R_{i,(z,x)}^2 = 1 - \frac{Var(u_{i,t})}{Var(dr_{i,t})}$$

which amounts to,

$$R_{i,(z,x)}^2 Var(dr_{i,t}) = Var(dr_{i,t}) - Var(u_{i,t})$$

showing that minimizing the residual variance over the set of possible pairs of factors is equivalent to maximizing  $R_{i,(z,x)}^2 Var(dr_{i,t})$ .

When solving this problem, weights may be applied to the maximized terms of interest rates at different maturities to capture the perceived relative importance of each one of them. That way, the problem becomes,

$$Max_{(z,x)} \sum_i \omega_i R_{i,(z,x)}^2 Var(dr_{i,t}) \quad (8)$$

The weights would help us to focus on the region of the TSIR relevant to a fund manager. If are using the technique to design a fixed income portfolio on public debt, we should weigh more heavily the maturities associated to bonds outstanding. Once again, the idea is that when dealing with a public or private debt market, we would have to start by estimating the TSIR, selecting a set of maturities, and generating interest rates time series for those maturities.

We also need to define *what do we understand by unexpected changes* in interest rates. Usually these are taken as differences between spot rates at  $t$  and the corresponding forward rate computed from the TSIR for time  $t-1$ , which agrees with the Pure Expectations Theory of the TSIR. Alternatively, the actual change in interest rates is used, as if it was completely unexpected. This corresponds to the assumption that each interest rates behaves as a random walk. For frequently observed data (weekly), differences between these two alternatives are small.

Once we have found the factors solving (8), we can estimate the sensitivity of  $r_i$  to changes in the factors, by estimating either (6) or (7). That way, for each interest rate in the original data set, we will have an estimate  $(a_i, b_i)$ , or an estimate  $(a_i - b_i, b_i)$ ,  $i = 1, 2, \dots$ .

From them, we can estimate sensitivity parameters for any other maturity by interpolating the estimated sensitivities, or by parameterizing a function like,

$$\begin{aligned} a_i &= c_0 + c_1 \ln i + c_2 (\ln i)^2 + c_3 (\ln i)^3 + \varepsilon_i, i = 1, 2, \dots, n \\ b_i &= d_0 + d_1 \ln i + d_2 (\ln i)^2 + d_3 (\ln i)^3 + \xi_i, i = 1, 2, \dots, n \end{aligned} \quad (9)$$

in the case of a one-factor model. Each one of these is a regression with a sample of  $n$  data points. Once we estimate these parameters, we can store them in a matrix:

$$B = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 \\ d_0 & d_1 & d_2 & d_3 \end{pmatrix}$$

When a one-factor model is considered, results on the optimal factor differ across countries. In Spain, the 3-year rate seems to do a good job [Navarro, E., and J.M. Nave (1997) "A two-factor duration model for interest rate risk management", *Investigaciones Económicas*]. For two-factor models, the 3-year and the 2-month rates were chosen in the Spanish market.

Naturally, *relating the factors used in TSIR in different markets*, and testing for the explanatory power of those from one market to the other, remains as an interesting issue for further research. Elton et al. found the six-year and the eight-month as the best factors. A standard result is that a third factor is usually needed to capture the curvature of the TSIR, since the resulting  $R$ -squared statistics are not very high for some maturities.

It is also interesting to point out that estimated sensitivities become sometimes non-significant in an interesting manner: in the Spanish market, for maturities over three years, interest rates are not sensitivity to the two-month rate factor, while for maturities below two-months, interest rate changes are not sensitive to changes in the three-year rate factor.

For the Spanish case, Navarro and Nave interpolate for the whole term structure using functions more general than the linear polynomials (9) above. Their functional specifications were,

$$\begin{aligned} (a - b)_i &= \frac{k_0 + k_1 \ln i + k_2 (\ln i)^2}{1 + k_3 \ln i + k_4 (\ln i)^2} \\ b_i &= \frac{h_0 + h_1 \ln i + h_2 (\ln i)^2}{1 + h_3 \ln i + h_4 (\ln i)^2 + h_5 (\ln i)^3} \end{aligned}$$

which should be constrained by,

$$\begin{aligned} (a - b)_{2\text{-month}} &= 1; \quad b_{3\text{-year}} = 1; \\ (a - b)_{3\text{-year}} &= 0; \quad b_{2\text{-month}} = 0; \end{aligned}$$

In addition, these authors also impose a condition on the derivatives of these functions,

$$(a - b)'_{3\text{-year}} = 0; \quad b'_{2\text{-month}} = 0;$$

The model obtained through this approach can be tested against competing alternatives (like one-factor models with different choices for proxies, or different two-factor models) by using estimated regressions of the type (6) or (7) to predict actual interest rate changes.

## 6.2 Portfolio management

Now, let us assume that we have to manage a fixed income portfolio by taking positions at time  $t$  in a given market, and present results at time  $t + h$ . We should then form expectations:  $E_t P_{t+h}$  for each bond trading in the market, and take short positions when we expect  $E_t P_{t+h} < P_t$ , and long positions when we expect the opposite. To be safe, we might take short positions only if we expect  $E_t P_{t+h} < (1 - \alpha)P_t$  and long positions only when we expect  $E_t P_{t+h} < (1 + \alpha)P_t$ , where  $\alpha$  might be 0.01 or 0.05, for instance. To form the expectation  $E_t P_{t+h}$  we need:

$$E_t P_{t+h} = E_t \sum_{i=1}^k C_{i,t+h+T_j} \delta_{t+h,T_j}$$

where  $C_{i,t+h+T_j}, j = 1, \dots, k$  denotes the  $k$  payments to be received before maturity, which will usually be known with certainty,  $T_j$  is the number of days after  $t + h$  when the payment will take place, and  $\delta_{t+h,T_j}$  is the appropriate discount factor to be applied at time  $t + h$ . Podemos considerar, como aproximación:

$$E_t \delta_{t+h,T_j} = (1 + E_t r_{t+h,T_j})^{-T_j/250}$$

or:

$$E_t \delta_{t+h,T_j} = \left( 1 - \frac{T_j}{250} E_t r_{t+h,T_j} \right)^{-1}$$

if  $T_j/250$  is less than one year, where  $E_t r_{t+h,T_j}$  is the expected (at time  $t$ ) interest rate to prevail in the market at time  $t + h$  for investments made at that time during  $T_j$  periods. If there is a large number of bonds trading in the market, the number of discount factors to be forecast can easily be too large. An advantage of the factor model is that it allows us to forecast all the discount factors from forecasts for just the two factors.

Using the regressions above, we estimate the implicit sensitivities to the two factors of the change in any interest rate at maturity  $T_j$  by:

$$\begin{pmatrix} \hat{a}_j \\ \hat{b}_j \end{pmatrix} = B \begin{pmatrix} 1 \\ \ln T_j \\ (\ln T_j)^2 \\ (\ln T_j)^3 \end{pmatrix} = \begin{pmatrix} c_0 + c_1 (\ln T_j) + c_2 (\ln T_j)^2 + c_3 (\ln T_j)^3 \\ d_0 + d_1 (\ln T_j) + d_2 (\ln T_j)^2 + d_3 (\ln T_j)^3 \end{pmatrix}$$



So, once we have available forecasts  $E_t dr_{x,t+h}$  and  $E_t dr_{z,t+h}$  for the variations between  $t$  and  $t+h$  of the changes in interest rates at maturities  $x$  and  $z$ , we can obtain a forecast for  $dr_{t+h,T_j}$  from:

$$E_t dr_{t+h,T_j} = (\hat{a}_j - \hat{b}_j) E_t dr_{z,t+h} + \hat{b}_j E_t dr_{x,t+h} \quad (10)$$

Once we have the forecast for  $dr_{t+h,T_j}$ , the change in  $r_{t+h,T_j}$ , we would compute the forecast for  $r_{t+h,T_j}$  itself, from:

$$E_t r_{t+h,T_j} = r_{t,T_j} + E_t dr_{t+h,T_j}$$

that reflects the fact that the expectation at time  $t$  of the interest rate prevailing at time  $t+h$  to be applied to investments for  $T_j$  days, is equal to the current (time- $t$ ) interest rate to be applied to investments for  $T_j$  days, plus our time- $t$  expectation of the change to be experienced by that discount factor between  $t$  and  $t+h$ .

In practice, we would generally start by doing this exercise for  $h=1$  period ahead, and we could then recursively move ahead to any future horizon  $h$ . So, we will have,

$$E_t r_{t+h,T_j} = r_{t,T_j} + \sum_{i=1}^h E_t dr_{t+i,T_j}$$

Hence, all we need is to start from some forecasts for the factors themselves:  $E_t dr_{z,t}$  and  $E_t dr_{x,t}$ . These might be obtained from univariate analysis of each time series of daily changes, or from a VAR model on both series of daily changes in interest rates. As an alternative to time series analysis, sometimes the analyst is interested on estimating the predicted change in the price of a bond under a given *scenario for interest rates*. Each scenario would be characterized by specific numerical values for  $dr_{z,t}$  and  $dr_{x,t}$ , and a given probability  $p$ . We could have several scenarios, with probabilities adding up to 1. Each scenario would be handled as if they were the forecasts coming from time series analysis.

Once we have forecasts for all the discount factors involved in a single bond, we could easily obtain a forecast for the price of the bond itself, and we might take a short or long position on that asset.

If we construct a matrix:

$$M = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \ln T_1 & \ln T_2 & \dots & \ln T_m \\ (\ln T_1)^2 & (\ln T_2)^2 & \dots & (\ln T_m)^2 \\ (\ln T_1)^3 & (\ln T_2)^3 & \dots & (\ln T_m)^3 \end{pmatrix}$$

where  $m$  is the (possibly large) number of dates in which a payment will be received from a given bond. The product  $B.M$  will give us a  $2 \times m$  matrix having the  $(a, b)$  coefficients for each maturity. Taken to (10) these coefficients will give us a forecast for the discount factor at each maturity:

$$E_t D_{t+h} = (B.M)' \begin{pmatrix} E_t dr_{z,t+h} \\ E_t dr_{x,t+h} - E_t dr_{z,t+h} \end{pmatrix}$$

with  $E_t D_{t+h} = (E_t \delta_{t+h+T_1}, \dots, E_t \delta_{t+h+T_m})$ . Finally, we can use these predictions to obtain the forecast for the price of the bond, which would help us decide whether we want to have that bond in our portfolio. If we want to forecast the market value of our fixed income portfolio, we would do a similar exercise for each bond in our portfolio.

### 6.3 A duration vector

The price of a bond paying coupon is,

$$P = \sum_{j=1}^k \frac{C}{(1+R_{t_j})^{t_j}} + \frac{N}{(1+R_{t_k})^{t_k}}$$

where sometimes a tax effect is modelled by,

$$P = \sum_{j=1}^k \frac{C(1+\gamma)}{(1+R_{t_j})^{t_j}} + \frac{N}{(1+R_{t_k})^{t_k}}$$

[see Vasicek, O.A. and Fong, H.G. (1982), "Term structure modeling using exponential splines", Journal of Finance].

From this expression,

$$dP = \sum_{j=1}^k \left[ \frac{\partial P}{\partial R_{t_j}} \frac{\partial R_{t_j}}{\partial R_{2\text{-month}}} dR_{2\text{-month}} + \frac{\partial P}{\partial R_{t_j}} \frac{\partial R_{t_j}}{\partial R_{3\text{-year}}} dR_{3\text{-year}} \right]$$

which can be seen to imply,

$$\frac{\Delta P}{P} = -D_1 \Delta R_{2\text{-month}} - D_2 \Delta R_{3\text{-year}} \quad (11)$$

where,

$$D_1 = \frac{\sum_{j=1}^k t_j (a-b)_{t_j} C (1+R_{t_j})^{-t_j-1} + t_k (a-b)_{t_k} N (1+R_{t_k})^{-t_k-1}}{P}$$

$$D_2 = \frac{\sum_{j=1}^k t_j b_{t_j} C (1+R_{t_j})^{-t_j-1} + t_k b_{t_k} N (1+R_{t_k})^{-t_k-1}}{P}$$

which become equal to Macaulay's modified duration in case the TSIR is flat and interest rate movements are of a parallel kind. This expression allows us to anticipate the percent price change in a bond that can be expected from changes in the two factors. That way a simulation of different scenarios can

be done, to estimate the change in the value of a bond that could be produced under each one of them.

The results can be extended to a portfolio of bonds. The relative price change caused on the portfolio by an interest rate movement can be estimated using a portfolio duration calculated as a weighted average of the duration of the bonds included in this portfolio, the weights being the percentage of the total portfolio assigned to each bond. It can be easily shown that an expression similar to (11) applies to changes in the market value of a portfolio, if durations  $D_1$  and  $D_2$  are defined as convex linear combinations of the analogue durations for the individual bonds, each one weighted by the relative proportion that each bond represents of total portfolio value.

Navarro and Nave provide an expression for the expected change in the market value of a portfolio under the Pure Expectations Theory of the term structure, that allows for testing for management quality in fixed income portfolios,

$$\Delta V = V(R)(1 + R_H)^H \left\{ \left[ \frac{H(a-b)_H}{1 + R_H} - D_1 \right] \Delta R_{2-month} + \left[ \frac{Hb_H}{1 + R_H} - D_2 \right] \Delta R_{3-year} \right\}$$

where  $H$  denotes the planning or investment horizon.

## 7 Componentes principales

El método de componentes principales que introducimos en esta sección permite reducir la dimensionalidad de un sistema de variables (rentabilidades financieras) mediante el uso de un conjunto relativamente reducido de combinaciones lineales de las mismas que sean capaces de explicar la mayor cantidad posible de variabilidad del vector de variables consideradas. La variabilidad de un vector se resume en su matriz de covarianzas, por lo que los componentes principales se construyen a partir de dicha matriz de covarianzas.

Pueden generarse tantos componentes principales como variables hay en el vector. El conjunto de todos los componentes principales contiene exactamente la misma información que el vector de variables original, pues es una transformación lineal del mismo. Sin embargo, como el interés estriba en reducir la dimensionalidad del mismo, se trata de explicar un porcentaje elevado de la variación temporal del vector utilizando un número reducido de componentes principales. Cuando el vector de variables original presenta fuertes correlaciones, entonces es generalmente posible encontrar que un número reducido de componentes principales explica un porcentaje elevado, digamos que del 90%, de la variabilidad global del conjunto de variables considerado.

Las combinaciones lineales que definen los componentes principales tienen como ponderaciones los autovectores asociados a los autovalores de la matriz de covarianzas del vector de variables. El autovector asociado al mayor autovalor define el primer componente principal, y así sucesivamente. La proporción de la varianza del vector explicada por cada componente principal viene dada por la relación existente entre el autovalor asociado y la suma de todos ellos. Al

tratarse de una matriz semidefinida positiva, los autovalores son todos no negativos, por lo que estas relaciones pueden interpretarse como una descomposición porcentual.

Los componentes principales son, por construcción, incorrelacionados. Ello puede interpretarse en el sentido de que la información que proporcionan con contiene elementos comunes. Se evita de este modo solapamientos que pudieran dar lugar a ineficiencias. Al tratar de explicar las fluctuaciones en el vector de variables por un número reducido de combinaciones lineales, parece interesante que cada una de ellas proporcione información distinta de la que aportan las demás; de este modo es más fácil que necesitemos un número pequeño de componentes principales.

La ausencia de correlación entre componentes principales es de gran utilidad para aproximar la matriz de covarianzas del vector original de variables, por ejemplo, en un análisis de simulación que tenga como objeto la valoración de una cartera de activos. Como la matriz de covarianzas de los componentes principales seleccionados es diagonal, es muy sencillo calcular a partir de ella la matriz de covarianzas del vector de variables original.

La ausencia de correlación cruzada entre los componentes principales es interesante para un gestor de carteras: supongamos que cada componente principal estimado para los activos de un determinado mercado nos ayuda a configurar una cartera de inversión en dicho mercado. Si las rentabilidades de dichas carteras estuvieran correlacionadas, tendríamos una inversión parcialmente redundante. Ello es ineficiente, pues, eliminando dichas redundancias podríamos reducir nuestra exposición en el mercado. Por el contrario, querríamos diseñar una o más carteras incorrelacionadas, que nos permitan cubrir al máximo la evolución del mercado, sin incurrir en las redundancias citadas.

Componentes principales es asimismo una técnica aplicable para estimar modelos en los que un elevado número de variables explicativas presenta alta colinealidad, o para tratar la ausencia de datos acerca de una variable cuando se dispone de información relativa a otras variables altamente correlacionadas con la primera.

## 7.1 El método de componentes principales

La técnica de Componentes Principales (CP) está diseñada para ser utilizada con variables estacionarias, por lo que se utiliza sobre rendimientos o sobre las variaciones diarias o mensuales de tipos de interés. Por tanto, podemos suponer que cada columna de  $X$  tiene media cero.

Partimos de una matriz de datos  $X$  de dimensión  $T \times k$  que tiene en columnas las observaciones correspondientes a cada una de las variables. El análisis de CP se basa en los autovalores y autovectores de la matriz  $V = X'X/T$ , de orden  $k \times k$ , que es la matriz de covarianzas de las variables. Cuando las variables consideradas tienen naturaleza diversa, conviene asimismo estandarizarlas pues de lo contrario, la estimación de los componentes principales puede estar dominada por las variables de mayor varianza. Si las variables se han estandarizado, su matriz de covarianzas será la matriz de correlaciones de las variables orig-

inales.<sup>10</sup> Calcular componentes principales sobre la matriz de covarianzas o sobre la matriz de correlaciones de las variables originales no da lugar a los mismos resultados, y no es siempre claro qué enfoque debe seguirse. En ocasiones, incluso puede ser útil seguir ambos enfoques y tratar de interpretar los resultados que se obtienen de ambos modos. En vectores de variables altamente correlacionadas, el primer CP, relacionado con el autovalor más grande será significativamente mayor que los demás autovalores, explicando un porcentaje elevado de la variabilidad conjunta.

Sea  $W$  la matriz  $k \times k$  que tiene por columnas los autovectores de  $V$ . Por tanto, tenemos la descomposición espectral de  $V$ :

$$V = W\Lambda W'$$

donde  $\Lambda$  es la matriz diagonal  $k \times k$  formada por los valores propios de  $V$ ,  $\lambda_i, i = 1, 2, \dots, k$ , y donde las columnas de  $W$  están ordenadas en consonancia con los elementos de la diagonal de  $\Lambda$ , en la que suponemos  $\lambda_1 > \lambda_2 > \dots > \lambda_k$ . Como la matriz de covarianzas (y la matriz de correlaciones) es positiva definida, todos sus autovalores son no negativos.

### 7.1.1 Autovectores normalizados

Si normalizamos cada autovector de modo que tenga norma unitaria (por ej., utilizando la norma euclídea), la matriz  $W$  es ortogonal:  $W^{-1} = W'$ , por lo que  $W'W = WW' = I_k$  y  $|W| = 1$ . Esto se debe a que los vectores propios correspondientes a dos valores propios distintos son ortogonales ente sí.

En consecuencia, el determinante de  $V$  es el producto de sus valores propios, y la traza de  $V$  es la suma de dichos autovalores:<sup>11</sup>

$$\begin{aligned} |V| &= |W\Lambda W'| = |W| |\Lambda| |W'| = |\Lambda| = \prod_{j=1}^k \lambda_j & (12) \\ trV &= tr(W\Lambda W') = tr(W'W\Lambda) = tr\Lambda = \sum_{j=1}^k \lambda_j \end{aligned}$$

El  $i$ -ésimo componente principal está definido por:

$$\begin{aligned} CP_i &= w_{1i}X_1 + w_{2i}X_2 + \dots + w_{ki}X_k = \underset{Txk \times kx1}{X} w_i \Rightarrow \\ &\Rightarrow CP_{it} = w_{1i}X_{1t} + w_{2i}X_{2t} + \dots + w_{ki}X_{kt} = X'_t w_i \end{aligned}$$

<sup>10</sup> Conviene examinar si el programa estadístico que se utiliza en el cálculo de los CPs trabaja con los valores y vectores propios de la matriz de covarianzas o de la matriz de correlaciones de los datos.

<sup>11</sup> Si trabajamos a partir de la matriz de correlaciones de las variables  $X$ , la matriz  $V$  tendría unos en su diagonal principal, por lo que la suma de los valores propios de  $V$  sería igual a  $k$ .

siendo una combinación lineal de las columnas de  $X$ , es decir, de las variables originales, con ponderaciones definidas por las coordenadas del  $i$ -ésimo vector propio de la matriz  $V$ . La matriz de todos los componentes principales es,

$$CP = \underset{Txk}{X} \underset{Txkxk}{W} \Rightarrow CP_i = Xw_{.i} = X \begin{pmatrix} w_{1i} \\ w_{2i} \\ \dots \\ w_{ki} \end{pmatrix} \quad (13)$$

A partir de la descomposición espectral tenemos,

$$VW = W\Lambda$$

por lo que,

$$(CP)'(CP) = W'X'XW = W'TVW = TW^{-1}W\Lambda = T\Lambda$$

que es una matriz diagonal, lo que prueba que:

- los componentes principales están incorrelacionados,
- si las variables originales  $X$  tienen media cero, los componentes principales también tienen media cero, y la varianza del CP  $i$ -ésimo es  $\lambda_i$ ,
- los componentes principales conservan la variabilidad de las variables originales: la suma de las varianzas de los CPs y de las variables originales son iguales:

$$Var(x_1) + \dots + Var(x_k) = tr(V) = \sum_{j=1}^k \lambda_j = \sum_{j=1}^k Var(CP_j)$$

- los componentes principales conservan el determinante de la matriz de covarianzas de las variables originales (su varianza generalizada). Por (12):

$$Var(CP) = Var(XW) = W'Var(X)W = W'VW = \Lambda \Rightarrow |Var(CP)| = |\Lambda|$$

- la proporción de la varianza del vector  $X$  que está explicada por el  $i$ -ésimo CP es  $\lambda_i / \sum_{j=1}^k \lambda_j$ . En efecto, puesto que  $\lambda_i$  es la varianza de  $CP_i$  y  $\sum_{j=1}^k \lambda_j = Var(x_1) + \dots + Var(x_k)$ .
- si las variables originales  $X$  tienen media cero, la covarianza entre cada componente principal y las variables originales viene dada por el producto del vector propio y el valor propio que definen dicho componente principal:

$$\begin{aligned}
Cov(CP, X) &= \frac{1}{T} (CP)' X = \frac{1}{T} (XW)' X = \frac{1}{T} W' (X' X) = W' V = \Lambda W' \Rightarrow \\
&\Rightarrow Cov(CP_i, X) = \frac{1}{T} (X w_{\cdot i})' X = \lambda_i w'_{\cdot i} = \lambda_i (w_{1i}, w_{2i}, \dots, w_{ki})'
\end{aligned}$$

- la correlación entre cada componente principal y las variables originales es proporcional al coeficiente de esa variable en la definición del componente principal. El coeficiente de proporcionalidad es el cociente entre la desviación típica del componente principal y la desviación típica de la variable:

$$Corr(CP_i, x_j) = \frac{Cov(CP_i, x_j)}{\sqrt{Var(CP_i)Var(x_j)}} = \frac{\lambda_i w_{ji}}{\sqrt{\lambda_i s_j^2}} = w_{ji} \frac{\sqrt{\lambda_i}}{\sqrt{s_j^2}}$$

Nótese que de acuerdo con este resultado, a igualdad de correlación entre dos variables y un determinado CP, tendrá mayor coeficiente en el mismo aquella variable que tenga mayor volatilidad. De nuevo, esta es una razón para estandarizar las variables  $X$  antes de aplicar la técnica de componentes principales

### 7.1.2 Elección del número de Componentes Principales

Una cuestión siempre importante es decidir el número de componentes principales que queremos elegir. Siempre será un criterio importante el porcentaje de varianza que queramos explicar, pero hay que tener en cuenta que esto dependerá de la naturaleza del problema. Alcanzaremos habitualmente elevados porcentajes de varianza trabajando en mercados de renta fija líquidos, y tendremos que contentarnos con porcentajes inferiores cuando trabajemos con mercados de renta variable. Por tanto, un primer criterio es tomar componentes principales hasta que el porcentaje de varianza explicado por el conjunto deje de aumentar apreciablemente.

Bai y Ng (2002) propusieron criterios que se han hecho muy populares. Denotemos por  $\hat{\sigma}^2(m)$  el promedio de las varianzas residuales en una regresión de cada una de las  $k$  variables sobre los primeros  $m$  factores:

$$\hat{\sigma}^2(m) = \frac{1}{k} \sum_{i=1}^k \hat{\sigma}_i^2(m)$$

Los criterios son:

$$\begin{aligned}
C_{p1}(m) &= \hat{\sigma}^2(m) + m \hat{\sigma}^2(M) \left( \frac{k+T}{kT} \right) \ln \left( \frac{kT}{k+T} \right) \\
C_{p2}(m) &= \hat{\sigma}^2(m) + m \hat{\sigma}^2(M) \left( \frac{k+T}{kT} \right) \ln (P_{kT}^2)
\end{aligned}$$

donde  $M$  es un número entero escogido previamente que denota el máximo número de factores, y  $P_{kT}^2 = \min(\sqrt{k}, \sqrt{T})$ . La idea es escoger  $m$  para minimizar  $C_{p1}(m)$  o  $C_{p2}(m)$ , con  $0 \leq m \leq M$ .

### 7.1.3 Representación de los datos en función de Componentes Principales

Puesto que  $W' = W^{-1}$ , la ecuación (13) es equivalente a:

$$X_{Txk} = (CP)_{Txk} W'_{kxk}$$

es decir,

$$X_i = w_{i1}CP_1 + w_{i2}CP_2 + \dots + w_{ik}CP_k$$

por lo que cada variable del sistema original puede escribirse como combinación lineal del conjunto de CPs, donde las ponderaciones son los elementos de la fila  $i$ -ésima de la matriz  $W$ . Esta es la *representación* del vector de variables en función de los CPs, y es la que facilita el cálculo de la matriz de covarianzas de  $X$ , puesto que, como hemos visto, la matriz de covarianzas de  $CP$  es  $\Lambda$ :

$$Var(X) = Var((CP) W') = W \cdot Var(CP) \cdot W' = W \Lambda W' = \sum_{i=1}^k \lambda_i (w_{.i} w'_{.i})$$

Esta representación nos sirve, de un lado, para tratar de interpretar con mayor precisión los componentes principales. Comparando el grado de ajuste de la regresión de cada tipo de interés sobre el primer componente principal, puede asociarse éste, a pesar de indicar el nivel general de tipos de interés, a un vencimiento concreto. De igual modo, estimando regresiones que utilizan el segundo componente principal para explicar distintos diferenciales entre tipos a largo y tipos a corto, puede asociarse el segundo componente principal con un determinado diferencial, como representación de la pendiente de la curva cupón cero. El tercer componente principal puede asociarse, igualmente por métodos de regresión, a la diferencia entre dos *spreads* de tipos de interés, generalmente, la diferencia entre un *spread* entre dos tipos a largo, y un *spread* entre dos tipos, ambos a corto plazo. Generalmente puede utilizarse un vencimiento intermedio como pivote en este cálculo.

Supongamos ahora que seleccionamos los  $n$  primeros componentes principales debido a que conjuntamente, explican un porcentaje suficientemente importante de la variabilidad del conjunto de variables originales,  $X$ . Es decir, supongamos que  $(\lambda_1 + \lambda_2 + \dots + \lambda_n) / \sum_{i=1}^k \lambda_i$  es del orden de 90% o incluso superior, y que  $n$  es pequeño: por ejemplo,  $n = 2$  ó  $n = 3$ . De hecho, por simplicidad, vamos a suponer  $n = 3$ . Tendríamos entonces:

$$X_i \simeq w_{i1}CP_1 + w_{i2}CP_2 + w_{i3}CP_3$$



$$Var(X) \simeq (w_{.1}, w_{.2}, w_{.3}) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} w'_{.1} \\ w'_{.2} \\ w'_{.3} \end{pmatrix} = \lambda_1 (w_{.1}w'_{.1}) + \lambda_2 (w_{.2}w'_{.2}) + \lambda_3 (w_{.3}w'_{.3})$$

y nos bastaría con tener predicciones de los tres componentes principales para un periodo futuro  $t + s$  para obtener predicciones de las variables en la matriz  $X$  para dicho periodo.

Si hemos trabajado con datos estandarizados, obtenidos a partir de los datos originales  $X^*$  mediante:<sup>12</sup>

$$X = (X^* - \bar{x}^* \circ 1_T) D^{-1}$$

donde  $D$  es una matriz diagonal que tiene como elementos las desviaciones típicas de las variables en  $X^*$ , tendremos:

$$(X^* - \bar{x}^* \circ 1_T) D^{-1} = (CP) W'$$

por lo que la representación final de las variables originales en función de los CPs es:

$$X^* = \bar{x}^* \circ 1_T + (CP) W' D$$

#### 7.1.4 Risk decomposition: Systematic and idiosyncratic risks

Suppose we need to estimate the covariance matrix of a large vector  $m \times 1$  of interest rates. The representation of each interest rate as a function of a *subset* of principal components can be written is,

$$r_{it}^* = w_{i1} cp_{1t} + \dots + w_{ik} cp_{kt}$$

with  $r_{it}^*$  being the standardized version of  $r_{it}$ ,  $\sigma_i = \sqrt{Var(r_i)}$ ,  $r_{it}^* = \frac{r_{it} - \bar{r}_i}{\sigma_i}$ .

Si seleccionamos un subconjunto de  $n$  componentes principales, then we have, in terms of original variables:

$$r_{it} = \bar{r}_i + w_{i1}^* cp_{1t} + \dots + w_{in}^* cp_{nt} + u_{it}$$

where  $w_{ij}^* = w_{ij} \sigma_j$  and where the error term indicates the approximation error when we take a subset of  $k$  of the  $m$  PCs.

Since PCs are uncorrelated, their variance-covariance matrix is,

$$V = W^* \Lambda W^{*'} + \hat{\Sigma}_u$$

where  $V$  is the variance-covariance matrix of the  $n$  interest rates,  $W^*$  is the  $m \times n$  matrix of weights,  $w_{ij}^*$ ,  $\Lambda$  is the  $n \times n$  diagonal matrix of standard deviations of the principal components, and  $\Sigma_u$  is the covariance matrix of approximation

<sup>12</sup>El producto  $\bar{x}^* \circ 1_T$  indica que el vector  $1_T$  queda multiplicado por cada elemento en  $\bar{x}^*$ , que es la media muestral de cada variable.

errors, which is not necessarily diagonal. The latter can be reduced by increasing the number of PCs considered in the representation of interest rates, and it is standard to ignore it in the previous expression.

This estimation of the variance-covariance matrix of interest rates may turn out to be positive semidefinite, if we take a number of PCs lower than the number of original variables. However, the big advantage is that we just need to estimate  $n$  variances, rather than  $m(m+1)/2$  variances and covariances.

In the decomposition above,  $W^* \Lambda W^{*'}$  represents the systematic risk in the set of assets, due to the uncertainty in the future evolution of the principal components, while  $\Sigma_u$  represents the size of idiosyncratic risk.

### 7.1.5 Componentes principales: una presentación alternativa

Un modo alternativo de considerar los CPs es el siguiente: Los CPs de un vector de rentabilidades  $r$  son combinaciones lineales del tipo:

$$y = rw$$

que tiene matriz de covarianzas:  $Var(y) = w' \Sigma_r w$ , con  $\Sigma_r = T^{-1} r r'$ . Queremos obtener un conjunto de combinaciones lineales incorrelacionadas, que tengan la mayor varianza posible. Pero esto podría conseguirse incrementando la norma de  $w$ . Como tal solución no tiene interés, nos restringimos a vectores  $w$  con  $w'w = 1$ , y consideramos:

$$\begin{aligned} &Max \ w_1' \Sigma_r w_1 \\ \text{suje}to \ a \quad &: \ w_1' w_1 = 1 \end{aligned}$$

con Lagrangiano:  $L(w_1, q) = w_1' \Sigma_r w_1 - q(w_1' w_1 - 1)$   
que tiene condiciones de optimalidad:

$$\begin{aligned} \frac{\partial L}{\partial w_1} &= 2 \Sigma_r w_1 - 2q w_1 = 0 \\ \frac{\partial L}{\partial q} &= w_1' w_1 - 1 = 0 \end{aligned}$$

por lo que  $w_1$  debe satisfacer:

$$\Sigma_r w_1 = q w_1$$

siendo por tanto un autovector de la matriz de covarianzas muestral  $\Sigma_r$ , mientras que el multiplicador de Lagrange es el autovalor asociado a dicho autovector.

Si multiplicamos por  $w_1'$  la ecuación anterior, tenemos:

$$w_1' \Sigma_r w_1 = q w_1' w_1 = q$$

por lo que  $w_1' \Sigma_r w_1$  se maximiza tomando como vector  $w_1$  en la combinación lineal el autovector asociado al mayor autovalor de  $\Sigma_r : q = \lambda_1$ .

El segundo CP será una combinación lineal de las rentabilidades en  $r$ , del tipo  $w_2' r$ , ortogonal (es decir, con covarianza y correlación lineal nulas) con el primer CP. Tenemos:

$$Cov(w_1' r, w_2' r) = w_1' \Sigma_r w_2 = T^{-1} w_1' r r' w_2$$

Si trasponemos, tenemos:

$$0 = T^{-1} (r r' w_2)' w_1 = (\Sigma_r w_2)' w_1$$

por lo que la covarianza entre ambas combinaciones lineales será igual a cero si  $w_2$  se elige de modo que se cumpla la relación anterior. Supongamos que tomamos como  $w_2$  un autovector de  $\Sigma_r$  distinto de  $w_1$ . Por ser autovector, su autovalor asociado  $\lambda$  satisface:  $\Sigma_r w_2 = \lambda w_2$ , y la ecuación anterior se convierte en:

$$0 = (\lambda w_2)' w_1 = \lambda w_2' w_1$$

que se cumple, por nuestra elección de  $w_2$ .

Ahora, si queremos resolver el problema:

$$\begin{aligned} & \text{Max } w_2' \Sigma_r w_2 \\ \text{su } & \text{jeto a } : w_2' w_2 = 1 \\ & w_2' w_1 = 0 \end{aligned}$$

lo que hemos de hacer es escoger como  $w_2$  un autovector de  $\Sigma_r$ , distinto de  $w_1$ , de modo que se maximice la forma cuadrática  $w_2' \Sigma_r w_2$ . Pero

$$w_2' \Sigma_r w_2 = w_2' (\lambda w_2) = \lambda w_2' w_2 = \lambda$$

que tomará un valor máximo si  $\lambda$  es el segundo mayor autovalor de la matriz de covarianzas (y correlaciones, debido a la estandarización)  $\Sigma_r$ .

Un argumento similar nos ayuda a construir el resto de los CPs del vector  $r$ .

**Example 9** *Let us suppose that we have two variables,  $y_1, y_2$ , which have been standardized to have zero mean and unit variance. Their covariance matrix  $S_y$  is then,*

$$S_y = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

with eigenvalues  $1 \pm \rho$ . If  $y_1, y_2$  are positively correlated, then the first (largest) eigenvalue is  $1 + \rho$ , with associated eigenvector  $(1, 1)$ . Hence, the first principal component of  $y_1, y_2$  is,

$$z_1 = \frac{1}{\sqrt{2}}(y_1 + y_2)$$

which is like an average of the original variables. Notice that, in fact, as we already know,  $Var(z_1) = 1 + \rho$ . The second principal component corresponds to the second (smallest) eigenvalue,  $1 - \rho$ , being,

$$z_2 = \frac{1}{\sqrt{2}}(y_1 - y_2)$$

which is proportional to the difference between the elements of  $y_1$  and  $y_2$ . It is easy to check that its variance is, in fact,  $Var(z_2) = 1 - \rho$ . If  $\rho < 0$ , the order of the eigenvectors would be reversed.

### 7.1.6 Lack of scale-invariance in principal components (Mardia, Kent, Bibby)

An important difficulty with principal components is that they are not scale-invariant, the reason being that eigenvectors are not scale-invariant either. If we divide each  $i$ -th variable by  $d_i$ , the covariance matrix of the new variables becomes  $DS_yD$ , where  $D = \text{diag}(d_i^{-1})$ . However, if  $x$  is an eigenvector of  $S_y$ , then  $D^{-1}x$  is *not* an eigenvector of  $DS_yD$ . That is, eigenvectors are not scale-invariant.

In the 2-variable case, we have,

$$S_y = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

where  $\rho > 0$ . The largest eigenvalue is  $\lambda_1 = \frac{1}{2}(\sigma_1^2 + \sigma_2^2) + \frac{1}{2}\Delta$ , with  $\Delta = [(\sigma_1^2 - \sigma_2^2)^2 + 4\sigma_1^2\sigma_2^2\rho^2]^{1/2}$ , and associated eigenvector proportional to,

$$(a_1, a_2) = [(\sigma_1^2 - \sigma_2^2) + \Delta, 2\rho\sigma_1\sigma_2] \quad (14)$$

When  $\sigma_1/\sigma_2 = 1$ , the ratio  $a_2/a_1 = 1$ . If  $\sigma_1 \neq \sigma_2$  and the first variable is multiplied by a factor  $k$ , we would like the  $a_2/a_1$ -ratio to also get multiplied by  $k$ . However, changing  $\sigma_1$  to  $k\sigma_1$  in (14) easily shows that this is not the case.

The lack of scale invariance means that we need to worry about the way scales are chosen and, in particular, about the units of measurement for each variable. Sometimes, variables are homogenized by standardizing them, i.e., dividing them by their respective sample standard deviation, finding principal components of their correlation matrix. However, this practice produces some difficulties in hypothesis testing.

### 7.1.7 Hypothesis testing on eigenvalues and eigenvectors

Under some regularity conditions, it can be shown that if  $l_1, l_2, \dots, l_k$  denote sample eigenvalues, then, asymptotically,

$$\ln l_j \sim N \left( \ln \lambda_j, \frac{2}{T-1} \right)$$

where the second argument is the variance of the distribution.

If  $g_j$  denotes the  $j$ -th sample eigenvector, and  $\gamma_j$  is the analogous population eigenvector, then we have asymptotically,

$$g_j \sim N_k \left( \gamma_j, \frac{1}{T-1} V_j \right)$$

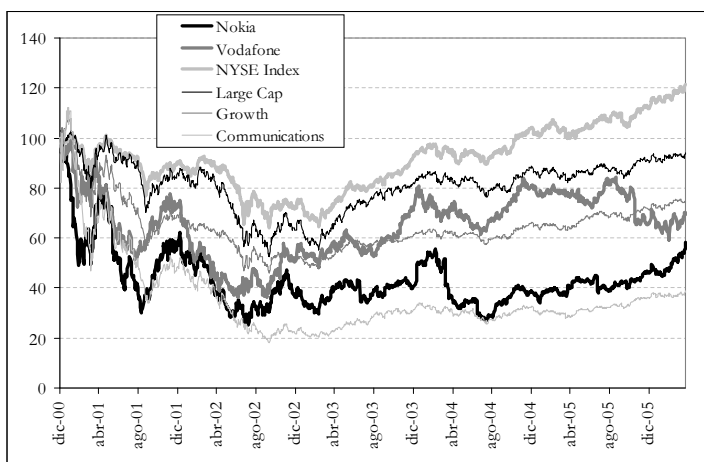
where,

$$V_j = \lambda_j \sum_{i \neq j} \frac{\lambda_i}{(\lambda_j - \lambda_i)^2} \gamma_j \gamma_i'$$

## 7.2 Algunas aplicaciones

### 7.2.1 Estimation of fundamental factor models in a stock market

The proposed *Case Study\_II.1* considers the risk decomposition of two stocks (Nokia and Vodafone) using historical prices and four fundamental risk factors: *i*) a broad market index, the NYSE composite index, *ii*) an industry factor, the Old Mutual communications fund, *iii*) a growth style factor, the Riverside growth fund, and *iv*) a capitalization factor, the AFBA Five Star Large Cap fund. The selection of the risk factors is a major issue, that depends on the user's experience and knowledge.



We consider a portfolio made up by 3 million US dollars of Nokia stock and 1 million US dollars of Vodafone stock. We start by estimating the total risk of the portfolio based on the historical returns on the two assets. The portfolio volatility is 42.5%. We estimate projections of each of the two assets on the four risk factors, to obtain the two vectors of 4 betas each. From that, we can

estimate the vector of portfolio betas. We also compute the covariance matrix of factor returns. The variance attributed to the risk factors, or *systematic variance*, is 24.7%, much lower than the total risk of the portfolio we estimated before. That means that the factor model does not explain the portfolio returns very well. Indeed, the R-squared statistics were 58.9% for Vodafone and 67.9% for Nokia. An additional problem is the important colinearity among the risk factors. The lowest correlation between any two of them is 0.69. Then, a set of regressions is estimated including one additional risk factor at a time, and the beta values can be seen to change dramatically.

As an approach to avoid the effect of collinearity, an orthogonal regression is also estimated for each asset using the first two principal components for the 4 risk factors. The systematic risk increases to 30.17%, but it remains still well below the estimate with historical portfolio returns.

### 7.3 Factor GARCH y GARCH ortogonal

Una vez que tenemos un modelo factorial:

$$r_{it} = \alpha_i + \beta_{i1}f_{1t} + \beta_{i2}f_{2t} + \dots + \beta_{im}f_{mt} + \varepsilon_{it}, \quad t = 1, 2, \dots, T; \quad i = 1, 2, \dots, k, \quad \text{con } m \ll k$$

la matriz de covarianzas de las rentabilidades  $r_{it}$  puede aproximarse:

$$\Sigma_r = B\Sigma_f B'$$

donde  $B$  es la matriz de las betas, de dimensión  $k \times m$  que tiene en cada fila las sensibilidades de cada activo a los  $m$  factores. La matriz  $\Sigma_f$  tiene dimensión  $m \times m$ . En el modelo Factor GARCH se permite que las varianzas y covarianzas de los factores evolucionen segun un modelo GARCH:

$$\Sigma_{r,t} = B\Sigma_{f,t}B'$$

Este podria ser un modelo DCC-GARCH, que es más sencillo de estimar, o un modelo GARCH multivariante si se trabaja con dos o tres factores. Lo que es importante es que la expresión anterior nos permite estimar matrices de covarianzas de gran dimensión a partir de una matriz de covarianzas de reducida dimensión, la correspondiente a los factores de riesgo. [ver ejercicio EII.4.7]

Un caso especialmente interesante resulta cuando se caracterizan factores mediante la técnica de componentes principales, en cuyo caso, la matriz  $\Sigma_f$  será diagonal. A pesar de ello, es fácil ver que la matriz  $\Sigma_r$  no lo será. Este modelo se conoce como orthogonal GARCH (O-GARCH). De hecho, este es un procedimiento que puede utilizarse para estimar correlaciones cambiantes en el tiempo, como alternativa al modelo DCC-GARCH. En efecto, la ecuación anterior nos proporcionará covarianzas y varianzas cambiantes en el tiempo, a partir de las cuales podremos estimar coeficientes de correlación.

### 7.3.1 Otros ejemplos de aplicaciones de la técnica de componentes principales (omitir)

- Tsay considera inicialmente rentabilidades mensuales de IBM, HP, Intel, Merrill Lynch y Morgan Stanley, de enero 1990 a diciembre 1999 (archivo m-5cln.txt). El modelo supone que las variables carecen de correlación serial. En caso contrario, hay que estimar previamente un modelo de series temporales apropiado para obtener unos residuos con estructura de ruido blanco a los que aplicar la estructura del modelo. Tras un análisis de CPs hecho previamente, se decide considerar 2 factores comunes, que se estiman por ambos procedimientos, que posteriormente se rotan por el criterio varimax que, en este caso, no introduce modificaciones sustanciales. Los factores resultantes recogen el sector financiero por un lado, y el sector tecnológico por el otro. Las varianzas específicas de IBM e Intel son elevadas, sugiriendo la existencia de aspectos específicos en estos valores que merecerían un estudio adicional.
- En un segundo ejemplo, Tsay considera las rentabilidades logarítmicas mensuales de índices de deuda a 1-, 5-, 10-, 20- y 30-años desde enero 1942 a diciembre 1999 (archivo m-bnd.txt). Se estima inicialmente un modelo VARMA(2,1) para eliminar la correlación serial, aunque ello no altera esencialmente la matriz de correlaciones de las rentabilidades. Ambos métodos de estimación muestran que 2 factores comunes explican más del 90% de la variación global. Las comunales son elevadas, y las varianzas específicas reducidas. El primer factor es un promedio de rentabilidades a distintos vencimientos, mientras que el segundo factor asigna ponderaciones correlacionadas positivamente con el plazo a vencimiento de cada índice. Los factores rotados asocian en el primer factor cargas crecientes con el tiempo a vencimiento, y decrecientes con el tiempo a vencimiento en el segundo factor.
- A continuación, se consideran las rentabilidades de 10 activos de renta variable (archivo m-excess-c10sp-9003.txt). Se comienza con un modelo de 2 factores que es rechazado por el contraste de Razón de Verosimilitudes. Se estima asimismo un modelo de 3 factores que no es rechazado por dicho contraste. Las correlaciones implicadas por el modelo de 3 factores estimado son más próximas a las muestrales que las que se obtienen del modelo de variables ficticias sectoriales que se había estimado en un ejemplo previo.
- Por último, se considera un ejemplo de lo que se conoce como análisis asintótico de Componentes Principales, correspondiente a una situación en que se cuenta con más variables que observaciones temporales. Para este caso, Connor y Korajczyk (1993) y Bai y Ng (2002) han propuesto criterios de selección del número de factores. Se consideran rentabilidades mensuales simples de 40 activos de renta variable entre enero 2001 y diciembre 2003, un total de 36 observaciones temporales. Los criterios citados sugieren

$m = 1$  y  $m = 6$  factores comunes, respectivamente. Se estima el modelo bajo la segunda opción, que explica un 89,4% de la variabilidad conjunta.

## 8 Applications of Permanent Components to Fixed Income management

### 8.1 Componentes principales en el análisis de una estructura temporal de tipos de interés.

The principal components technique is designed to find factors explaining most of the variance in a vector of time series. The factors are then linear combinations of the original variables. The principal components technique consists of computing the eigenvalues of the variance matrix of the standardized variables. The eigenvectors associated to the eigenvalues (they are all positive) of highest size define the linear combinations to be used as principal components. It can be shown that the proportion of variance in the original data set that each principal component is able to explain is given by the size of the associated eigenvalue as a proportion of the sum of all them. An advantage of the technique is that it is very simple to implement. A drawback is that the obtained factors may not easy be interpret. However, when we can find an interpretation for them, the analysis may gain in interest.

When working with data from a given fixed income market, we start by identifying a vector of interest rates to summarize through their principal components. To do so, we might have to start by estimating a term structure. Once the term structure has been estimated, we may select a vector of maturities, and the associates interest rates form the vector of time series to analyze. For reasons of nonstationarity of interest rates, it is useful to start by computing correlations among changes in interest rates.

In many international fixed income markets three factors are enough to explain most of the variance across the TSIR.<sup>13</sup> Furthermore, their interpretation is the same in all cases: the first factor captures the general level of interest rates. Changes in this component can be seen as parallel shifts in the TSIR. The second component captures the slope of the TS, while the third component represents the curvature of the TSIR, and describes changes in the concavity/convexity of the TSIR. This result has direct a bearing on immunization, suggesting that a standard approach of choosing a portfolio having as duration the investment horizon produces a less than complete immunization, since it covers only the risk involved in parallel shifts of the TSIR. Hence, the level of risk covered can be estimated by the proportion of variance explained by the

---

<sup>13</sup>D'Ecclesia, L. and S.A. Zenios (1994) "Risk factor analysis and portfolio immunization in the Italian bond market", The Journal of Fixed Income, sept., p. 51-58, Navarro, E. and J.M. Nave (1995) "Análisis de los factores de riesgo en el mercado español de deuda pública", Cuadernos Aragonese de Economía, 5, 2, 331-341, Steeley, J.M. (1990), "Modelling the dynamics of the term structure of interest rates", The Economic and Social review, 21, 4, 337-361.



first principal component. The implication then is that a technique of vector immunization, covering against multiple risk factors, should lead to improved results. One of the possibilities that has been explored is to construct functions using the time series of estimated parameters in an interest rate model like the one proposed by Nelson and Siegel.

Una alternativa al análisis de regresión que antes vimos para la caracterización de los factores de riesgo de una estructura temporal consiste en calcular los componentes principales de un amplio vector de tipos de interés cupón cero. En la mayoría de los casos, el primer componente principal tiene sus coordenadas de magnitud aproximadamente similar, generalmente con pesos ligeramente menores en los vencimiento extremos. Se interpreta, por tanto, como un promedio de todos los tipos de interés considerados, por lo que viene a indicar el nivel general de tipos de interés. Por tanto, las variaciones imprevistas en el nivel general de tipos de interés constituyen el factor más importante de fluctuaciones en la estructura temporal.

La serie temporal que puede construirse para este primer componente principal utilizando las ponderaciones que lo definen indicaría la evolución temporal histórica del nivel general de tipos de interés en la economía. Su correlación con cualquiera de los tipos de interés inicialmente considerados será elevada, si bien estaremos tratando generalmente con variables integradas, y dichas correlaciones no pueden interpretarse de modo muy estricto.

El segundo componente principal suele estar formado asignando ponderaciones de signo opuesto a los tipos de interés en los vencimientos cortos y largos. Además, el valor absoluto de dichas ponderaciones suele cambiar suavemente, desde un extremo de la estructura temporal al otro. Se interpreta como la pendiente de la estructura temporal, lo que sugiere que variaciones inesperadas en la pendiente de la curva es la segunda causa más importante de fluctuaciones en el vector de tipos de interés.

El tercer componente principal suele asignar ponderaciones de igual signo a ambos extremos del rango de vencimientos, y ponderaciones de signo opuesto a las primeras al tramo central de vencimientos. Sus variaciones pueden interpretarse, por tanto, como cambios en la curvatura de la estructura temporal, siendo estos la tercera causa, en importancia, de fluctuaciones en la estructura temporal de tipos de interés.

**Example 10** *Example 6.2.1, Alexander(2001) analiza la estructura temporal para tipos cupón cero en EEUU, con datos mensuales de enero 1944 a diciembre 1992. Son tipos de interés semianualizados, tomados de Coleman, T.S., L. Fisher y R. Ibbotson (1992), "U.S. Treasury Yield Curves 1926-1992", Ibbotson Associates, Chicago. El primer autovector explica 78,6% de la varianza de los 14 tipos considerados. Los dos primeros CPs explican 90,3%, y los tres primeros explican un 93,8%. Por razones de estacionariedad, se trabaja con variaciones mensuales en los tipos de interés, usando su matriz de correlaciones. La interpretación de los CPs es similar a la apuntada antes.*

**Example 11** *Example II.2.8, Alexander considera la representación por CPs de dos estructuras temporales simultáneamente. Para ello, trabaja con 10 vencimien-*

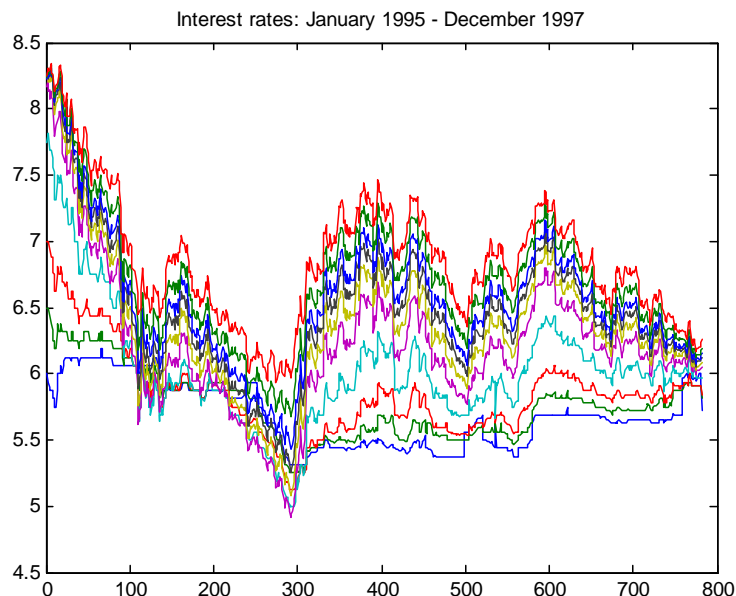
tos de la curva de EEUU y 11 vencimientos de la curva del Reino Unido, seleccionando 4 CPs. Los dos primeros CPs explican, respectivamente, un 71,3% y un 24,7% de la variación total en el vector de 21 tipos de interés. El primero de ellos es un promedio global de tipos mientras el segundo parece ser el diferencial de pendientes en ambos países. Tendría interés examinar la relación entre los CPs que pueden caracterizarse analizando las curvas de cada país por separado, es decir, el grado en que las variaciones en los CPs de la curva de tipos de un país pueden explicar variaciones en la curva de tipos de otro país.

**Example 12** *Un análisis similar, aplicado a las pendientes de la estructura temporal puede verse en "A factor model of term structure slopes in eurocurrency markets", Applied Economics Letters, (2002), 9,5 85-593, E.Dominguez y A.Novales, as well as in "An Error Correction Factor Model of Term Structure Slopes in International Swap Markets", Journal of International Financial Markets Institutions & Money, P.Abad y A. Novales. Dado el contenido informativo que la pendiente tiene para anticipar el ciclo económico en muchos países, es interesante analizar si los cambios en la pendiente de la curva de tipos presenta aspectos comunes para distintos países.*

**Example 13** *Alexander (2001) considera posteriormente la estructura temporal de contratos NYMEX de futuros sobre petróleo sin refinar, con vencimientos de 1 a 12 meses, en el período de 4 de febrero de 1993 a 24 de marzo de 1999. El primer CP explica ya un 96% de la variación global en las rentabilidades, y puede asociarse con desplazamientos paralelos en todos los precios. El segundo CP, con interpretación de pendiente de la curva, explica un 3% adicional.*

## 8.2 Principal components analysis of a set of interest rates

Consider the set of interest rates contained in *US\_tipos.prn*. This is daily data from January 5, 1995, to December 31, 1997 on interest rates at maturities: 1-, 3-, 6- and 12-months, and 2-, 3-, 4-, 5-, 7- and 10-years. When the original data is a set of variables similar in nature, as in this case, the need to transform them into stationarity or to standardize them is not so important. So, we are going to analyze the data without any transformation in spite of the fact that the interest rates are not stationary:

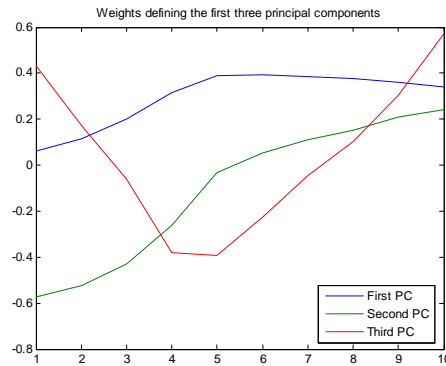


As we can see, interest rates at different maturities tend to move together, but they also exhibit notorious differences. The slope of the term structure remained positive along the sample period (the red line is the 10-year interest rate, and the blue line at the bottom of the graph is the 1-month rate). The slope narrowed initially and then rapidly increased, to decrease again towards the end of the sample period. The variances of the different interest rates are: 0.0596; 0.0667; 0.1027; 0.2029; 0.2863; 0.2914; 0.2809; 0.2693; 0.2535; 0.2344. The sum of all them is 2.0478. The eigenvalues of the variance covariance matrix of interest rates are: 0.0000, 0.0000, 0.0001, 0.0002, 0.0003, 0.0009, 0.0031, 0.0184, 0.1448, 1.8799, that also add up to 2.0478. It is not very good that some of these eigenvalues are essentially zero. That says that some of the interest rates are almost redundant, i.e., they can be approximately generated as linear combinations of other interest rates in the vector. This is not strictly true, but there is so much correlation among the ten interest rates, that the variance-covariance matrix becomes singular. The singularity might not be so obvious if we had transformed the interest rates as suggested at the beginning of this section.

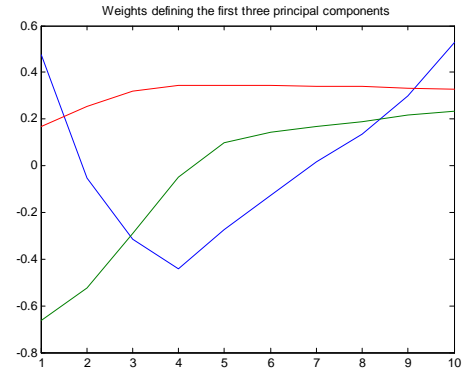
Therefore, the first principal component explains 91,8% of the variance of the vector interest rates, the first two together explain 98,9% of that variance, and the first three explain 99,8%. We will retain just these first three principal components. They are defined by the weights given by: PC1: (0.0629, 0.1169, 0.1997, 0.3175, 0.3879, 0.3926, 0.3851, 0.3759, 0.3614, 0.3418) in blue; PC2: (-0.5735, -0.5254, -0.4315, -0.2603, -0.0332, 0.0550, 0.111, 0.1518, 0.2078, 0.2414) in green; PC3: (0.4297, 0.1733, -0.0601, -0.3816, -0.3917, -0.2251, -0.0463, 0.1041

0.3024 0.5724) in red (left graph):

Non-standardized rates



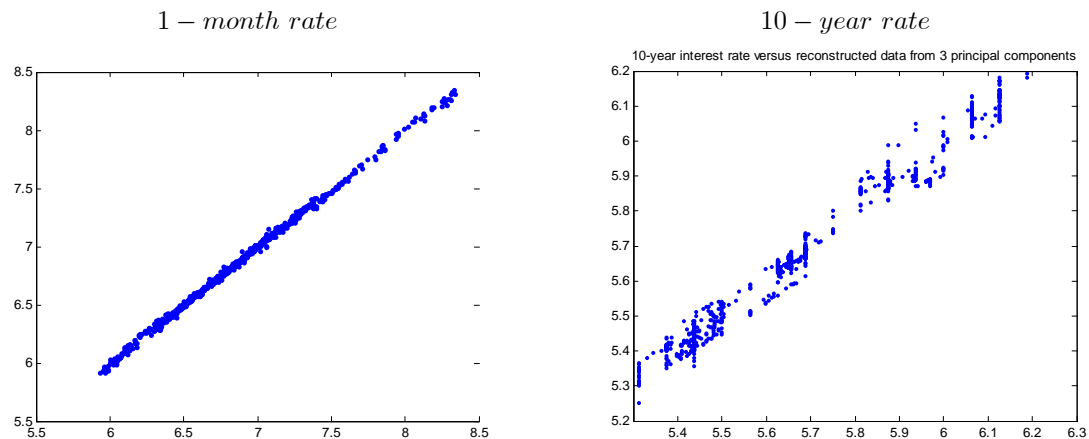
Standardized interest rates



that have the standard interpretation of capturing the general level of interest rates, the slope of the term structure, and its curvature.

If we work with standardized interest rates, the eigenvalues are: ( 0.0001, 0.0001, 0.0003, 0.0008, 0.0030, 0.0069, 0.0255, 0.1061, 1.6868, 8.1703), so that none of them is zero, even though some are small, relative to the largest one. The first component explains 81,7% of the fluctuation in the vector interest rate, the first two components explain 98,6%, and the first three explain 99,5% of that joint variance. These three components can be characterized by their weights, shown in the right graph above, which are not all that different from the ones we obtained without standardizing the original variables. Now we could proceed with the same procedure we explained above.

Once we limit our attention to a few principal components, three in this case, we can recover the data from them as a good approximation, since we are missing just 0,2% of the information in the original data set. The left graph shows that the comparison between the shortest maturity interest rate (1-month) and the approximation we would get from the first three principal components looks quite good while the right graph shows that the ability to replicate interest rates at the longest maturities is more limited.



This is due, in part, to the fact that there were many days in the sample in which the 7-year and the 10-year interest rate did not change. There is not much we can do about that. These approximations have been obtained multiplying the  $T \times 10$  data matrix by the transpose of the *last three columns* of  $W$  (since in Matlab eigenvalues and eigenvectors are ordered from smallest to largest), transposed to form a  $3 \times 10$  matrix:

$$\hat{X}_{T \times 10} = (CP_8 CP_9 CP_{10})_{T \times 3} (W_8 W_9 W_{10})'_{10 \times 3} \quad (15)$$

The last three columns of  $W$  give us the 3 Betas for each interest rate as a function of the three chosen principal components. As a matter of fact, the same approximation to the original data as we have described from the product matrix can be obtained by least squares, estimating a regression of each interest rate on the three chosen principal components. The fitted time series from the regression and the time series recovered from the matrix product above will have a correlation of almost 1.0.

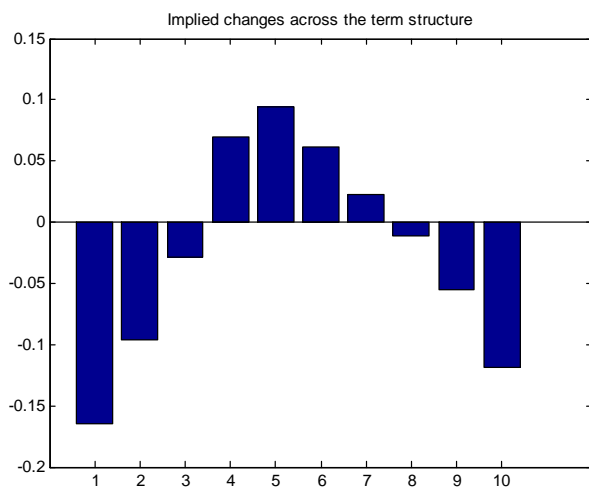
There are many uses of the principal component representation. Let us describe the main one, regarding portfolio/risk management in fixed income market. Suppose we have a portfolio in such market with 50 bonds in it. We know the current market value of the portfolio  $P_T$ , but we want to figure out the possible evolution of that price in the horizon of  $h$  days, that is, between  $T$  and  $T+h$ . To do so, we must have an idea about the possible values of the discount function at  $T+h$ . We will need to discount all the cash flows remaining at  $T+h$  and that can be a number significantly higher than the number of bonds. Suppose that at  $T+h$  there will be 300 cash flows to be paid on the portfolio bonds. The maturities will easily be arbitrary: 13 days, 56 days, 86 days, 143 days, and so on. We assume that the cash flow paying the farthest away is less than 10 years.

We start from the interest rate database that we have analyzed and for which we have already characterized three principal components as explained above.

We assume these were zero-coupon interest rates. Now, we have to figure out what evolution these principal components may follow between  $T$  and  $T + h$ . Once we do that, representation (15) will allow us to get the expected evolution for the 10 interest rates, by doing:

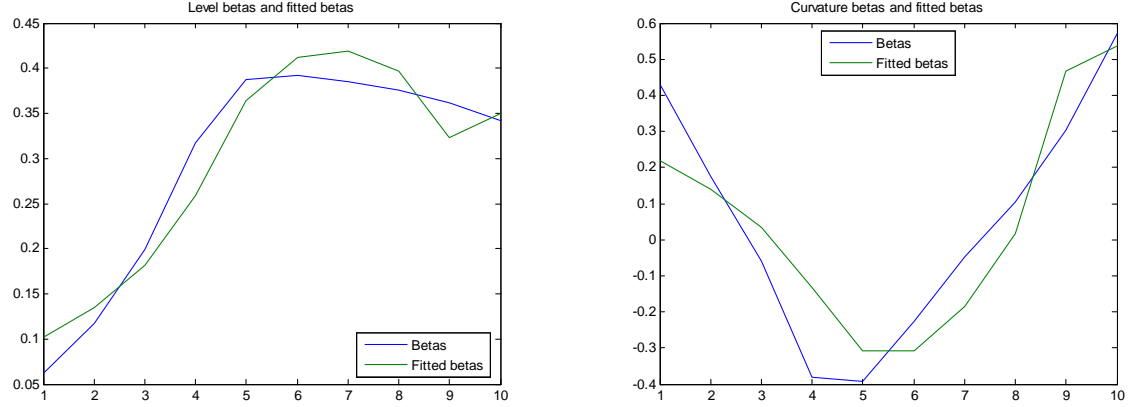
$$\Delta X = (\Delta CP_8 \Delta CP_9 \Delta CP_{10})(W_8 W_9 W_{10})'$$

Suppose we consider a first scenario in which the first principal component decreases by 25 basis points as a consequence of a cut in the intervention interest rate by the monetary authority, with an increase in slope of 10 basis points,<sup>14</sup> with no change in curvature. Suppose that that is our maintained scenario for changes in the permanent components between  $T$  and  $T + h$ . The implied changes for the different maturities would be:



We have solved the first stage of the analysis. To get a model that can be used to forecast prices for a given fixed income portfolio, we follow the approach we have described in previous sections. At this point, we have an scenario at  $T + h$  for interest rates at fixed-maturities, but what we really need is an scenario for the discount function. One possibility would be to interpolate on the previous graph over the maturities, to have the predicted change in the discount function at each possible maturity. But this could be computationally complex. It might be easier to do the following. We fit a polynomial to each of the principal component betas. For each principal component, we have a beta for each interest rate, for a total of 10 betas. That is not a lot of data, but we can fit a third order polynomial in log-maturity, for instance:

<sup>14</sup>This change in slope may happen because the short term rate comes down by 25 basis points while the longest end of the term structure decreases by only 15 basis points.



with an analogous curve for slope betas. Notice that the horizontal axis describes the maturities from 0 to 10-years. So, if we are given a 734 days cash flow, say, we can find the betas for that maturity. All we need to do is to use the estimated regression coefficients to form a  $4 \times 3$  matrix  $\beta$ :

$$\beta = \begin{pmatrix} & \textit{Level} & \textit{Slope} & \textit{Curvature} \\ \text{Constant} & 0.0840 & -0.5942 & 0.2602 \\ m & 0.2151 & 0.3589 & -0.5198 \\ m^2 & -0.0422 & -0.0546 & 0.1338 \\ m^3 & 0.0023 & 0.0027 & -0.0079 \end{pmatrix} \quad (16)$$

and the product  $\beta_m = \beta^t M$  where  $M = (1, m, m^2, m^3)'$  which is  $4 \times 300$ , will give us the 3-vector of betas  $\beta_m$  for maturity  $m$ . All we need to do is to keep in memory the vector of 300 maturities, from which we can immediately construct a  $4 \times 300$  matrix  $M$  generalizing the  $M$ -vector above, and premultiplying by  $\beta^t$ , we have a  $3 \times 300$  matrix of betas, a 3-vector for each of the 300 maturities. That way we would obtain a value for each zero coupon rate at each maturity, and the discount factor is obtained as usual:  $d_m = (1 + r_m)^{-1}$ .

The final stage relates to constructing the scenarios for  $T + h$ . There are essentially two possibilities: *i*) to write down a number of  $r$  scenarios, similar to the one we described above. They would try to capture the different alternatives for monetary policy or for market fluctuations we consider reasonable for the next  $h$  market days. Next, we have to assign a probability to each scenario, reflecting the likelihood we associate to each one. This could be by assigning a weight to each scenario and dividing each weight by their sum. That way, we get something looking like probabilities. Since each scenario will be defined in terms of a given change in level, in slope and in curvature of the term structure, we could follow the steps above to translate each scenario into an implied change for the vector of 300 discount factors.

The alternative procedure consists of *ii*) modelling the time series evolution of each fixed-maturity interest rate. This could be made with univariate processes or in a multivariate fashion. Working with a multivariate process for

10 variables is going to be rather complex, but we can simplify it by specifying processes describing the time evolution of the three principal components. Not only we have a shorter number of variables, but also, since they are uncorrelated, we do not gain anything by fitting multivariate model. Hence, we can fit a univariate model to each principal component, and then run a large number of simulations, 5000 say, for each one of them. That would give us an empirical density for each principal component at time  $T + h$ . Suppose we take the estimated mean value for each one of those densities at  $T + h$ . The principal component betas would give us the implied values for the 10 interest rates, and the  $\beta$ -matrix representation above (16) would give us the values for the 200 discount factors. We could have taken the median values for the components at  $T + h$ . These empirical densities obtained by simulation would allow us to also compute an estimate of VaR for our portfolio by taking some extreme percentiles of each density in the direction of a decrease in the portfolio market price.

Some more observations:

- Stationarity would have suggested to work with daily differences in interest rates. The first principal component then explains only 62.9% of the joint variance in interest rate changes. This is a consequence of daily changes being far more volatile than interest rate levels. The first two components explain 86,7% of the variance in daily changes in interest rates, and the first three components explain 94,6%.
- A completely different approach could consist of estimating a given analytical model for the zero coupon curve, like the Nelson\_Siegel model. If we estimate the model with daily data for a given period, we would have time series for each of the parameters in the model, and a multivariate model (there are 4 parameters in the Nelson-Siegel model) could help us to simulate their behavior or to predict their values at the desired horizon. Using the predicted values (or the mean values for each parameter from its empirical density if we have a run a large number of simulations) we could compute the whole yield curve at the desired horizon. Then, we can read on that curve directly the zero coupon discount rate and hence, the discount factor associated to each maturity at which a bond is paying a coupon after  $T + h$ . In this case, we would be taking as risk factors the parameters in the Nelson-Siegel model. But they will not be uncorrelated, so care must be paid when running the simulations to preserve the correlation structure observed in actual data.
- Predicción de tipos de interés utilizando Componentes Principales: Una cuestión sin duda importante se refiere a la posibilidad de utilizar la capacidad explicativa de los componentes principales con fines *predictivos*. Indudablemente, esta es una cuestión de la que, en principio, no cabría esperar una mejoría sobre el uso de las variables originales, por cuanto que los componentes principales son combinaciones lineales de valores contemporáneos de las variables, pero facilitaría enormemente la tarea de predecir



un alto número de tipos de interés. El enfoque trataría de utilizar regresiones estimadas de cada variable (tipo de interés) sobre los dos o tres primeros componentes principales, para predecir, es decir, lo que hemos denominado la representación del vector original de variables en función de los CPs seleccionados. Como los CPs se estiman con datos contemporáneos del vector de variables considerado, tal ejercicio de predicción requiere previamente la predicción de los propios componentes, agregando así errores de predicción. Aunque esto podría deteriorar la calidad de las predicciones, es sorprendente observar que la capacidad predictiva de un modelo como el citado es, en ocasiones, muy buena. Una regresión estática de un número reducido de componentes principales puede predecir al menos tan bien como el modelo univariante de cada tipo de interés [A modo de ejemplo, ver Domínguez y Novales (2002)].

### 8.3 Present value of a basis point: PV01

A preliminary concept is  $\delta 01$ , the *basis point sensitivity of the discount factor*, defined as the *increase* in the discount factor because of a *fall* of 1 basis point in the relevant interest rate,

$$\delta 01_t = \frac{1}{(1 + R_t - .01)^t} - \frac{1}{(1 + R_t)^t}$$

where we initially assume  $T$  to be an integer number of years. The basis point sensitivity,  $\delta 01$ , is usually less than 10 basis points, it increases with maturity and the increase is more pronounced for low levels of interest rates. It is always a *positive* number.

For a sequence of cash-flows, the *present value of a basis point* is defined as:

$$PV01(C, r) = PV(C, r^-) - PV(C, r) > 0$$

where  $PV(C, r)$  denotes the present value of the cash flow sequence  $C$  (a  $k \times 1$  vector), and  $r$  is the discount rate vector, while  $r^- = r - (.01\%)1_k$ , measures the absolute change in the value of a cash flow sequence because of a *fall* of one basis point in the yield curve (term structure). Esto es lo que Hull (*Risk Management and Financial Institutions*, 2007) denomina DV01: the change in the value of a portfolio for a one-basis-point parallel shift in the yield curve.

For a single payment  $C_t$  at time  $t$ , the *present value of a basis point* can be obtained by multiplying the money amount by the change in the discount factor as a consequence of the change of one basis point in interest rates at that maturity:

$$PV01(C_t) = C_t \cdot \delta 01_t > 0$$

and for a sequence of cash flows over time,

$$PV01(C_{t_1}, C_{t_2}, \dots, C_{t_k}) = \sum_{i=1}^k C_{t_i} \cdot \delta 01_{t_i}$$

Para un bono determinado, PV01 aproxima la diferencia en el valor presente de la secuencia de cash flows y, por tanto, el cambio en el precio del bono. For a single bond, PV01 is very similar to the standard duration, which is the change in present value of the bond per unit change in the bond yield. There are two differences:

1. In general, the PV01 is the exact cash flow sensitivity to a parallel shift in the zero coupon yield curve, whereas duration is just an approximation. Both concepts coincide just if a shift in the zero coupon curve caused the yields on all the bonds in the portfolio to change by the same amount, which is very unlikely.
2. Duration cannot be extended to cover all interest rate sensitive instruments. For instance, it cannot apply to negative cash flows.

**Example 14** *Example III.1.18: A bond with nominal 1 million pounds, that pays a 6% coupon until maturity, 4 years from now, trades at 103.62, exactly one year from the next coupon. Current zero coupon rates at 1, 2, 3 and 4 years are 4.50%, 4.75%, 4.85% and 5.00%, respectively. Compute the present value of a basis point for this bond and compare it with Macaulay's duration. Solution: The PV01 is 36,312.75, which will be the profit made by the holder if zero coupon rates decrease by one basis point at all maturities. The increase in price will be from the current 103.6216 to 103.6579. The yield (internal rate of return) of the bond is 4.9796%, and the value duration, which approximates the change in the price of the bond, is 36,326.59.*

## 8.4 Approximations to PV01

Let us derive a first order (linear) approximation to  $\delta 01$  and hence, from that, an approximation to PV01. If  $t$  is an integer number of years, the discount factor is  $\delta_t = (1+R_t)^{-t}$  and so, for a single basis point change ( $\Delta R = -0,01\% = -0,0001$ ), we have:

$$\delta 01_t \simeq -\frac{d\delta_t}{dR_t} 10^{-4} = t(1+R_t)^{-(t+1)} 10^{-4}$$

If there is a change of  $m$  basis points, this expression must be multiplied by  $m$ . The Present Value of a given cash-flow at time  $t$  is:  $PV(C_t, R_t) = C_t(1+R_t)^{-t} = C_t\delta_t$ .

Then, a useful approximation to the PV01 of a cash flow  $C_t$  at time  $t$  when  $t$  is an integral number of years is:

$$PV01(C_t) = \frac{\partial PV(C_t, R_t)}{\partial R_t} = C_t\delta 01_t \approx tC_t(1+R_t)^{-(t+1)} 10^{-4}$$

When  $t$  is less than a year,

$$\delta_t = (1+tR_t)^{-1} \Rightarrow \delta 01_t \simeq t(1+tR_t)^{-2} 10^{-4}$$

Sometimes, continuously compounded interest rates are preferred. The continuously compounded discount factor is,

$$\delta_t = \exp(-r_t t)$$

En términos de tipos continuos, si  $r_t$  es el tipo de interés que compuesto continuamente, genera el mismo tipo de interés discreto  $R_t$ , tenemos:  $PV(C_t, R_t) = C_t \exp(-r_t t)$ , with  $r_t$  being the continuously compounded rate of interest at maturity  $t$ .

Thus, we have approximate expressions for  $PV01$  valid for any maturity.

$$\delta 01_t \simeq -\frac{d\delta_t}{dr_t} 10^{-4} = t \exp(-r_t t) 10^{-4} = t \delta_t 10^{-4}$$

and therefore, the approximation to the  $PV01$  for a cash flow at any maturity under continuous compounding is:

$$PV01(C_t) \approx t C_t \exp(-r_t t) 10^{-4}$$

**Example 15** Example III.1.19: Compare the exact and the approximate values for  $PV01$  in a cash flow that has been mapped to vertices 1 and 2 years, with 10 million euros mapped to the 1-year vertex and 5 million euros mapped to the 2-year vertex. Suppose current zero coupon rates at 1 and 2 years are 4.00% and 4.50%, respectively.  $R$ : The exact  $PV01$ , calculated from the change of a 1 basis point in the discount factor is 1801,07 euros, while the approximate  $PV01$ , calculated from the expression introduced in this section, is 1801,85 euros.

## 8.5 Riesgo en carteras de renta fija

Consideremos ahora una cartera de renta fija. Puede tratarse de una larga relación de flujos de caja, a pagar o cobrar en el futuro, o puede tratarse de un bono o de una cartera de bonos.

Si hemos caracterizado los factores de riesgo mediante Componentes Principales, recordemos que si  $\delta 01_t$  denota la sensibilidad del factor descuento aplicable al vencimiento  $t$  ante una variación de un punto básico en el tipo cupón cero a dicho plazo, tenemos:  $PV01(C_t) = C_t \cdot \delta 01_t$  que depende, por supuesto, del tamaño del flujo de caja a dicho plazo.

La variación en el precio de la cartera de renta fija es la variación en el valor presente de la secuencia de cash-flows, multiplicados por la magnitud de la variación en cada tipo de interés:

$$P_t - P_{t-1} = -\sum_{i=1}^k C_{t_i} \cdot \delta 01_{t_i} (R_{i,t} - R_{i,t-1}) = -\sum_{i=1}^k PV01(C_{t_i}) (R_{i,t} - R_{i,t-1}) = -\sum_{i=1}^k PV01(C_{t_i}) \Delta R_{i,t}$$

o en notación matricial:

$$\Delta P_t = -\theta' \Delta R_t$$

siendo  $\theta$  el vector formado por los  $PV01$  de los distintos flujos de caja:  $\theta = (PV01_1, PV01_2, \dots, PV01_k) > 0$ .

Si utilizamos  $n$  componentes principales  $CP$  como factores de riesgo, la varianza de las variaciones diarias en tipos de interés cupón cero puede escribirse, en función de dichos factores:

$$Var(\Delta R_t) = Var_{Txn \times nxk}(CP^* W^{*'}) = W^* Var(CP^*) W^{*'} = W^* \Lambda^* W^{*'}$$

donde  $\Lambda^*$  denota la matriz diagonal formada con los autovalores asociados a los componentes principales seleccionados  $CP^*$ , y  $W^*$  las columnas formadas por los autovectores correspondientes. Si hemos escogido 3 Componentes Principales tendríamos para cada tipo de interés:

$$Var(\Delta R_{it}) = \sum_{j=1}^3 w_{ij}^2 \lambda_j$$

Por tanto, para el valor de mercado de la cartera tendríamos:

$$Var(\Delta P_t) = \theta' Var(\Delta R_t) \theta = (W^{*'} \theta)' \Lambda^* (W^{*'} \theta) = S' \Lambda^* S$$

donde  $S = W^{*'} \theta$ , es un vector  $nx1$ , contiene las sensibilidades que habitualmente calculamos para una secuencia de flujos de caja respecto de los  $n$  factores que hemos seleccionado. Las hemos calculado multiplicando el vector de  $VP01$  por las coordenadas que definen cada uno de los componentes principales. Dada la estructura diagonal de la matriz de covarianzas  $\Lambda^*$  de los Componentes Principales, tenemos finalmente:

$$Var(\Delta P_t) = \sum_{j=1}^3 S_j^2 \lambda_j$$

siendo  $S_j$  la sensibilidad de la secuencia de flujos de caja respecto de cada componente principal.

## 8.6 Interest rate risk

Let us think of today as time zero:  $t = 0$ . Today, we can use current zero coupon rates to compute the current present value of a cash flow  $C_T$ , to be paid at time  $T > 0$ . There is no uncertainty involved in that. We also know for certain what the present value will be at maturity. But the present value at any intermediate time  $t$  between today ( $t = 0$ ) and maturity,  $t = T$ , is uncertain today because of uncertainty on:

- by how much the discount rate might change, and
- how sensitive is the present value to changes in the discount rate

To answer the first question we need a measure of variance for the discount rate. To answer the second, we use the concept of present value of one basis point.

It is usually assumed that the best forecast of the appropriate discount rate at time  $t$  is the forward zero coupon rate  $F_{t,T-t}$ , starting at time  $t^*$  with maturity  $T - t^*$ . We calculate forward rates today using current interest rates. Under such assumption, today's expected discounted value of  $C_T$  at time  $t^*$  is  $C_T$  times the appropriate discount factor, based on this forward rate:

$$\begin{aligned} PV_{t^*}(C_T) &= \frac{C_T}{(1 + F_{t^*,T-t^*})^{-(T-t^*)}} \\ PV_{t^*}(C_T) &= C_T e^{-(T-t^*)f_{t^*,T-t^*}} \end{aligned}$$

The sensitivity of this time  $t$  value to movements in the forward rate is measured by the  $PV01$ . Assuming both are integer numbers, we have the approximation:

$$\begin{aligned} PV01_{t^*}(C_T) &\approx -\frac{\delta(PV_{t^*}(C_T))}{\delta F_{t^*,T-t^*}} 10^{-4} = (T-t^*)C_T(1 + F_{t^*,T-t^*})^{-(T-t^*+1)} 10^{-4} \\ PV01_{t^*}(C_T) &\approx -\frac{\delta(PV_{t^*}(C_T))}{\delta f_{t^*,T-t^*}} 10^{-4} = (T-t^*)C_T e^{-(T-t^*)f_{t^*,T-t^*}} 10^{-4} \end{aligned}$$

The interest rate risk of a cash-flow of  $C_T$  refers to the uncertainty about the discounted value of this cash flow at some future time  $t^*$  between  $t$  (today) and  $T$  (maturity). Estamos considerando el riesgo inherente a las fluctuaciones que pueda experimentar desde hoy hasta  $t^*$  el factor de descuento (el tipo forward) que haya que aplicar a  $C_T$ . En  $t^*$  dicho factor de descuento será el tipo spot cupón cero entre  $t^*$  y  $T$ . Assuming that changes in the forward rate are independent and identically distributed, the standard deviation of the forward rate over  $(T - t^*)$  periods can be obtained by using the *square-root rule*:  $\sqrt{(T - t^*)}\sigma_{F_{t^*,T-t^*}}$ .

The standard deviation of the future discounted value above is the product of both quantities:

$$\sigma(PV_{t^*}(C_T)) \approx 10^{-4} \sqrt{(T - t^*)}\sigma_{F_{t^*,T-t^*}} (T - t^*)C_T(1 + F_{t^*,T-t^*})^{-(T-t^*+1)}$$

With continuously compounded forward rates:

$$\sigma(PV_{t^*}(C_T)) = 10^{-4} \sqrt{(T - t^*)}\sigma_{f_{t^*,T-t^*}} (T - t^*)C_T \exp[-(T - t^*)f_{t^*,T-t^*}]$$

[ExIII.1.20]

La percepción del riesgo de tipos de interés depende de si el interés es fijo o es flotante y también dependen de los criterios contables. En el trading book,

los instrumentos de renta fija entran con su valor mark-to-market. Por tanto, su riesgo depende de la evolución de los factores de descuento del modo que acabamos de ver. Por el contrario, en el banking book los instrumentos de renta fija entran según el principio de devengo, no según su mark-to-market. Los cash-flows (flujos de caja) se valoran en el momento del pago, de modo que los instrumentos de renta fija no están sujetos a riesgo de tipo de interés, aunque sí a otros tipos de riesgo, como riesgo de contraparte. Los instrumentos de tipo de interés flotante no tienen riesgo en el mark-to-market, pero sí en la contabilidad según el principio de devengo, porque el flujo de caja que se vaya a recibir en una fecha de pago futura es incierto.

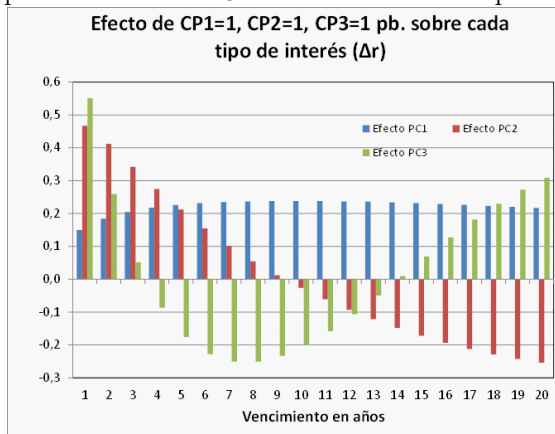
### 8.7 Summary of expressions

	Discrete compounding	Continuous compounding
Discount factor	$\delta_T = (1 + R_T)^{-T}$	$\delta_T = \exp(-r_T T)$
Interest rate sensitivity	$\delta 01_T \simeq T(1 + R_T)^{-(T+1)} 10^{-4}$	$\delta 01_T \simeq T \exp(-r_T T) 10^{-4}$
PV of one basis point for $C_T$	$PV01_T \approx TC_T(1 + R_T)^{-(T+1)} 10^{-4}$	$PV01_T \approx TC_T \exp(-r_T T) 10^{-4}$
Expected PV of $C_T$ at $t$	$PV01_{t,T} \approx (T - t)C_T(1 + F_{t,T-t})^{-(T-t+1)} 10^{-4}$	$PV01_{t,T} \approx (T - t)C_T \exp[-(r_{t,T-t})(T-t)] 10^{-4}$

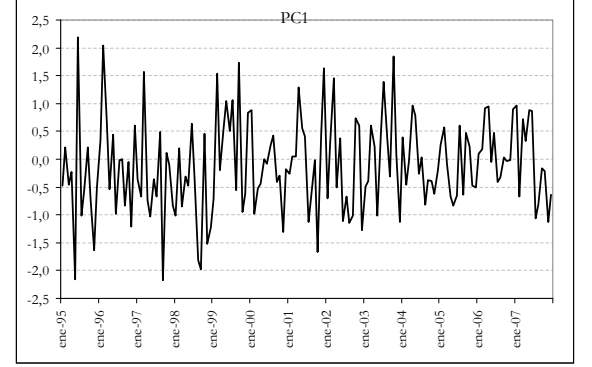
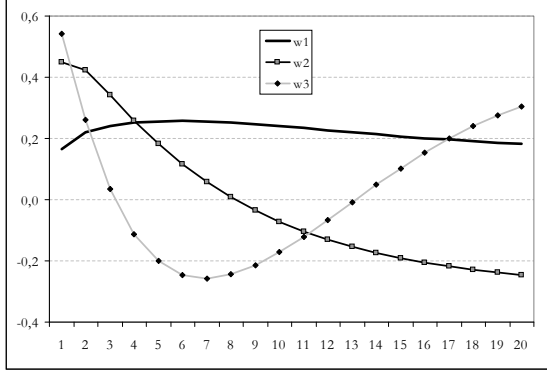
### 8.8 Varias aplicaciones

**Example 16** Example II.2.1: *Principal Component factor model for a UK bond portfolio* (<http://www.bankofengland.co.uk/statistics/yieldcurve/index.htm>).

A portfolio of UK government bonds has been mapped to interest rates at maturities 1 year, 2 years, ..., 20 years. The cash flow over 20 years is given. We are supposed to build a PC factor model for this portfolio.



The first three principal components explain more than 99% of the joint fluctuation in the term structure.



Remember that the change in price of a bond can be approximated by the change in interest rate at each relevant maturity, multiplied by the present value of one basis point at each of those maturities:

$$\Delta P_t = -\theta' \Delta R_t$$

where  $\theta \equiv PV01$  denotes the vector made up by the present value of a one basis point at each maturity, which is a function of the size of the cash flow at that maturity. The vector  $\theta \equiv PV01$  is positive by construction, measuring the effect of a *decrease* in interest rates on the present value of a given cash-flow. When interest rates rise, present values go down, and so does the price of any bond, which explains the minus sign in the expression above.

But once we construct the Principal Components, we can approximate the change in each interest rate as a function of changes in the factors (the principal components in this case):

$$\Delta R_{i,t} = w_{i1}PC_{1t} + w_{i2}PC_{2t} + \dots + w_{in}PC_{nt}$$

where  $(w_{i1}, w_{i2}, \dots, w_{in})$  is the vector of loadings of the interest rate change  $\Delta R_{i,t}$  (the *i*th-row of matrix  $W$ ).

Then,

$$\Delta P_t = - \sum_{i=1}^k [PV01(C_{ti})] \cdot \Delta R_{it} = - \sum_{i=1}^k PV01(C_{ti}) \left( \sum_{j=1}^n w_{ij} PC_{jt} \right) = - \sum_{j=1}^n \left( \sum_{i=1}^k PV01(C_{ti}) w_{ij} \right) PC_{jt} = - \sum_{j=1}^n \beta_j PC_{jt} \quad (17)$$

with the beta on each Principal Component being defined by,

$$\beta_j = \sum_{i=1}^k PV01(C_{t_i}) w_{ij} = \sum_{i=1}^k C_{t_i} \cdot \delta 01_{t_i} w_{ij}, \quad j = 1, 2, \dots, k \quad (18)$$

Coming back to the example, we can use this expression to compute the sensitivities of the Profit&Loss of the portfolio to the PCA factors:

$$P\&L_t \simeq -\mathbb{E}623.74PC_{1t} - \mathbb{E}1001.11PC_{2t} + \mathbb{E}1006.65PC_{3t}$$

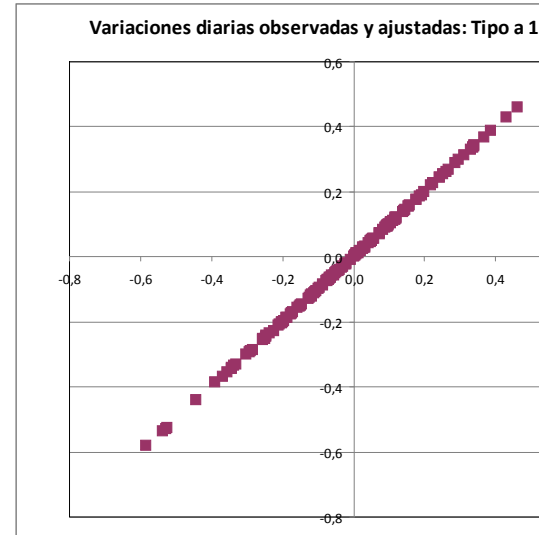
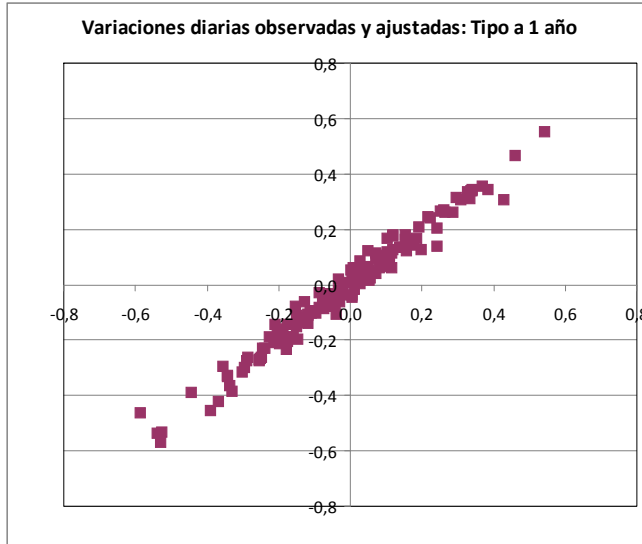
that provides us with an estimate of the approximate change in price as a change of any given magnitude in each Principal Component. Since the PCs can generally be interpreted as the general level in interest rates, the slope of the term structure and its curvature, this expression allows us to compute the expected change in the nominal value of the portfolio of alternative.

For instance, consider a reduction of 1% (100 bps.) in all interest rates produced by a change in just the first  $PC$ . Since the weights on the first  $PC$  are very similar, with an average of 0,22, then for each interest rate:  $\Delta R_{i,t} = w_{i1}PC_{1t}$ , we should have  $PC_{it} \simeq -100/(0,22) = -454,55$ , so that the first term effect on  $P\&L_t$  would be:  $(-623,74)(-454,55) = \mathbb{E}283518,2$ , a figure similar to the one obtained in the 'PV01' tab of EXII.2.1.

If we compute the principal components out of the correlation matrix of daily changes in interest rates, the Profit & Loss approximation would be:

$$P\&L_t \simeq -\mathbb{E}473.51PC_{1t} - \mathbb{E}1143.70PC_{2t} + \mathbb{E}992.69PC_{3t}$$

with an estimated gain of  $(-473,51)(-454,55) = \mathbb{E}215234,0$ . We can use the representation expression to approximate the observed interest rates as a function of the risk factors. We will alternatively choose three and five principal components as such risk factors. For the one-year rate, the relation between the observed one-year rate and the approximations are:



**Example 17** Example II.2.2: *Factor model for currency forward positions:*

At time  $t$  we face a sequence of foreign currency payments  $(C_1, \dots, C_n)$  at future times  $(T_1, \dots, T_n)$ . Let us denote by  $\Delta P_t^d$  the change in the present value



of the entire sequence of cash flows in domestic currency when the domestic interest rates change by amounts:

$$\Delta R_t^d = (\Delta R_{1t}^d, \dots, \Delta R_{nt}^d)$$

Then,

$$\Delta P_t^d = - \sum_{i=1}^n PV01_i^d \Delta R_{it}^d$$

where  $PV01_i^d$  is the  $PV01$  sensitivity of the cash flow in domestic currency at maturity  $T_i$ . Similarly,

$$\Delta P_t^f = - \sum_{i=1}^n PV01_i^f \Delta R_{it}^d \quad (19)$$

is the change in present value of the sequence of payments in foreign currency when the domestic interest rates change. Here, only the cash-flows are in foreign currency, interest rates are domestic.

If  $S_t$  denotes the domestic foreign exchange rate, then:  $P_t^d = S_t P_t^f$ , so that taking logs and differencing over time, we get:  $R_t^d \simeq R_t^s + R_t^f$ , where  $R_t^d$  is the return on the cash flow in domestic currency,  $R_t^f$  is the return on the cash flow in foreign currency, and  $R_t^s$  is the return on the spot exchange rate.

Now we can approximately decompose the risk on a sequence of foreign currency forward payments into exchange rate and interest rate risks:

$$Var(R_t^d) \simeq Var(R_t^s) + Var(R_t^f) + 2Cov(R_t^s, R_t^f)$$

However, although the exchange rate risk is defined in terms of the variance of returns, the interest rate risk is defined in terms of the variance of the  $P\&L$  position, and not the variance of returns. But we can rewrite the previous expression:

$$Var\left(\frac{\Delta P_t^d}{P_{t-1}^d}\right) \simeq Var(R_t^s) + Var\left(\frac{\Delta P_t^f}{P_{t-1}^f}\right) + 2Cov\left(R_t^s, \frac{\Delta P_t^f}{P_{t-1}^f}\right)$$

Now, we approximate  $P_{t-1}^d \simeq \bar{P}^d$ ,  $P_{t-1}^f \simeq \bar{P}^f$ , where  $\bar{P}^f, \bar{P}^d$  denote the present values of the cash flows in domestic and foreign currencies respectively,

$$Var\left(\frac{\Delta P_t^d}{\bar{P}^d}\right) \simeq Var(R_t^s) + Var\left(\frac{\Delta P_t^f}{\bar{P}^f}\right) + 2Cov\left(R_t^s, \frac{\Delta P_t^f}{\bar{P}^f}\right)$$

and we multiply through by  $(\bar{P}^d)^2$ , to get:

$$Var(\Delta P_t^d) = (\bar{P}^d)^2 Var(R_t^s) + \bar{S}^2 Var(\Delta P_t^f) + 2\bar{P}^d \bar{S} Cov(R_t^s, \Delta P_t^f)$$

where  $\bar{S}$  is the exchange rate at the time the risk is measured. Thus  $\bar{P}^f, \bar{P}^d$  and  $\bar{S}$  are known at time  $t$ . The terms on the right hand side are quadratic forms based on covariance matrices of a very large number of different domestic interest rates. In the next example, we consider a schedule of 60 monthly foreign currency payments, so  $Var(\Delta P_t^f)$  would be calculated from a quadratic form with a  $60 \times 60$  covariance matrix, while  $Cov(R_t^s, \Delta P_t^f)$  would have 60 components. These would clearly be very costly to compute. But, assuming that the currency is very liquid, forward prices will be very close to their fair value, and Principal Components turn out to be very helpful:

1. We start by using the term structure of interest rates from each country to compute the present value of the sequence of cash-flows. If we used the exchange rate to write the two present values in the domestic currency they would not be equal to each other, because we have used a different term structure to compute them.
2. The first term is easy: we compute the volatility of returns in the exchange rate in our sample, and multiply it by the price of the portfolio (sequence of cash flows) at the domestic currency.
3. For the second term, we want to compute the variance of changes in the price of the portfolio at the foreign currency. Remember that:  $\Delta P_t^f = -\sum_{i=1}^k PV01_i^f \Delta R_{it}^d = -\sum_{j=1}^n \beta_j PC_{jt}$ , with  $\beta_j = \sum_{i=1}^k PV01_i^f w_{ij}$ ,  $j = 1, 2, \dots, n$ . If we denote by  $\lambda_j$  the eigenvalues of the covariance/correlation matrix of interest rates:  $Var(\Delta P_t^f) = \sum_{j=1}^n \beta_j^2 Var(PC_{jt}) = \sum_{j=1}^n \beta_j^2 \lambda_j$ .
4. For the last term, we need the Covariance between daily returns on the exchange rate and portfolio price changes in the foreign currency,  $Cov(R_t^s, \Delta P_t^f)$ . We will use the representation of  $\Delta P_t^f$  as a function of the Principal Components:  $\Delta P_t^f = -\sum_{j=1}^n \beta_j PC_{jt}$ , to get:  $Cov(R_t^s, \Delta P_t^f) = -\sum_{j=1}^n \beta_j Cov(R_t^s, PC_{jt})$ . So, we first compute the covariance between the time series for each Principal Component and the returns on the exchange rate. Each of the three covariances is multiplied by the corresponding portfolio beta,  $\beta_j = \sum_{i=1}^k PV01_i^f w_{ij}$ , and we aggregate these products. Finally, that covariance is multiplied by the exchange rate and the present value of the cash flow sequence in the domestic currency:  

$$Correlation\ risk = \sqrt{2\bar{P}^d \bar{S} Cov(R_t^s, \Delta P_t^f)}$$

**Example 18** *Let us suppose that a UK refinery has purchased crude oil in the US paying in US dollars through futures contracts, so that there is no commodity price risk. However, some other risks remain: i) exchange rate risk, arising from uncertainty on the sterling value of future payments in dollars, ii) interest rate risk, arising from the change in present value of the sterling cash flows, iii) the correlation risk, arising from the correlation between UK interest rates and the sterling-dollar exchange rate (see exchange rate data and interest rate data in*

<http://www.bankofengland.co.uk/>). We assume the UK company has forward payments of \$1 million on the 5th of every month over the next 5 years.

We first approximate the change in present value of the sequence of payments in foreign currency when the domestic interest rates change:  $\Delta P_t^{\$} = 1788.45PC_{1t} - 412.71PC_{2t} + 331.68PC_{3t}$  and use this expression and the variances of the three Principal Components to obtain:  $Var(\Delta P_t^{\$}) = (43, 479, 68)^2$ . Then, we estimate the interest rate component of risk:  $IRrisk = \sqrt{250} \bar{S} \text{€} = \sqrt{250}(43, 479, 68)/(1, 9909) = \text{€}345.308$ . The 250 factor is used because of dealing with daily data.

The volatility of the exchange rate is obtained as  $\sqrt{250Var(R_t^s)}$ , and the foreign exchange risk is obtained multiplying the volatility of the exchange rate by the present value of the cash-flow in the domestic currency:  $FXrisk = \sqrt{(27.067.101)^2 250Var(R_t^s)} = (27.067.011) (7.83\%) = \text{€}2.118.347$ . For Correlation risk, we use expression  $Cov(R_t^s, \Delta P_t^f) = \sum_{j=1}^k \beta_j Cov(R_t^s, PC_{jt})$ , introducing again a 250 factor. The result is 5.211, 62. After that, we multiply the covariance estimate by the present value of the sequence of cash flows in the domestic currency and by the exchange rate.  $Correlation-risk = \sqrt{2(250)Cov(R_t^s, \Delta P_t^f)}(27.067.101)/(1, 9909) = \text{€}376.442$ . Finally,  $TOTAL RISK$  is equal to  $\text{€}2.179.069$ , most of it explained by exchange rate risk.

**Example 19** Example II.2.3: *PCA factor models for crude oil (commodity) futures*

**Example 20** Example II.2.4: *Bond portfolio immunization. Use the same data from Example II.2.1. Let us now see how to immunize this portfolio against the main sources of risk in the term structure of interest rates. We assume the risk factors are appropriately summarized by the first three principal components. We now ask the question of how much of the 10-year bond should we add to our portfolio so that the new portfolio Profit&Loss is invariant to changes in the first principal component.*

The present value of the cash flow sequence to be immunized is  $\text{€}9.410.205, 2$ . We take this as an approximation to the immunization of the sequence of cash-flows against parallel shifts in the term structure. As shown in the spreadsheet, we need a negative cash flow of  $\text{€}-4.225.241, 7$  at 10-year maturity, which amounts to taking a short position of  $\text{€}-2.716.823, 5$ , that is, selling now this nominal amount of the 10-year bond. In fact, it is possible that such an exact amount cannot be sold, and the financial manager will have to set up for an approximation. Immunization would require trading for about half of the present value of the cash flow sequence to be immunized.

With that strategy, we would get the factor risk model:

$$P\&L_t \simeq \text{€}1190, 88PC_{2t} - \text{€}558, 54PC_{3t}$$

The choice of maturity is arbitrary, so long as there is a bond instrument that can be traded.

Now we could ask which positions should we take in the 5- and 15-year bonds to immunize the cash-flow sequence against changes in slope of the term structure. We will need to take positions at two different maturities to achieve that hedge. The answer is a negative cash flow of  $\text{€}-9.745.454,5$  at the 5-year maturity, and a long position of  $\text{€}2.202.790,5$  at 15-year maturity, which amount to selling now  $\text{€}-7.852.637,2$  of the 5-year bond, and purchasing  $\text{€}1.142.974,3$  of the 15-year bond. With these added positions, the risk factor model becomes,

$$P\&L_t \simeq -\text{€}82,14PC_{3t}$$

and we could additionally find three positions at different maturities that could immunize the cash flow sequence against changes in the three principal components simultaneously.

**Example 21** Example II.2.5: *Asset-liability management: A single curve PCA factor model can be used to balance assets and liabilities. For example, consider a pension fund that asks how to invest its income from contributors in fixed income securities so that its P&L is insensitive to the most common movements in interest rates, as captured by the first three principal components.*

Similarly, a corporate may have a series of fixed liabilities, such as payments on a fixed rate loan, and seek to finance these payments by issuing fixed coupon bonds or notes. In Example II.2.5, we consider a UK firm with a fixed stream of liabilities of £1 million per month over the next 5 years. It seeks to finance these by issuing zero coupon bonds at 1-month, 1-, 3- and 5-years to maturity. How many bonds should it issue or purchase on 31 December 2007 so that the portfolio of assets and liabilities has zero sensitivity to parallel shifts and changes in slope of the UK government spot yield curve?

The present value of the liabilities on 31 December 2007 is £53.887.892. The factor model for the original stream of cash flows is,

$$P\&L_t \simeq \text{€}1768,18PC_{1t} - \text{€}406,33PC_{2t} + \text{€}325,01PC_{3t}$$

We want to find cash flows at 1-, 3- and 5-year maturities such that a) the present value of the increased sequence of cash flows remains the same as before, £53.887.892, and b) the net position of assets and liabilities has a Profit&Loss that has zero sensitivities to the first and second principal components of the UK spot curve. As shown in the spreadsheet, if we add cash flows of £4.257.977,3; £16.952.130,9; £29.308.732,0 and £9.483.229,8 at the 1-month, 1-, 3- and 5-year maturities, respectively, the new stream of cash flows will have the required zero sensitivities.

**Example 22** Example II.2.6: *Stress testing a UK bond portfolio. Stress tests have become unfortunately popular during the current financial crisis. They consist of passing a given portfolio through what is considered an extreme market situation. Of course, the definition of what we mean by such situation is crucial for the result of the test. We start this exercise by computing the portfolio Profit&Loss volatility based on a one-, two-, and three-principal component*

representation of the term structure, and comparing the results with the ones obtained without using the factor model.

From the Profit&Loss factor representation,

$$P\&L_t \simeq \text{€}623.74PC_{1t} + \text{€}1001.11PC_{2t} - \text{€}1006.65PC_{3t}$$

we get an estimate of the variance of the  $P\&L_t$ , based on one principal component, of  $Volatility(P\&L_t) = \sqrt{12\beta_1^2 Var(PC_1)} = \text{€}178,897$ . If we base the computation on two components, we get:  $Volatility(P\&L_t) = \sqrt{12 [\beta_1^2 Var(PC_1) + \beta_2^2 Var(PC_2)]} = \text{€}207,639$ , while if we use three principal components we get:  $Volatility(P\&L_t) = \sqrt{12 [\beta_1^2 Var(PC_1) + \beta_2^2 Var(PC_2) + \beta_3^2 Var(PC_3)]} = \text{€}211,214$ . The variance necessarily increases by increasing the number of principal components we use in this estimation, but the result we obtain with two and three components is very similar.

The alternative procedure consists of using the representation based on the present value of a one basis point change:  $\Delta P_t = \sum_{i=1}^n PV01_i \Delta R_{it} = PV01' \Delta R_t$ , to obtain:

$$Volatility(P\&L_t) = \sqrt{12 (PV01) [Var(\Delta R_{it})] (PV01)'} = \text{€}211,463$$

very close to the result obtained with the principal components. Since we are dealing with monthly data, a factor of 12 was used to compute the  $Var(P\&L_t)$ .

We now move to the proper stress testing, by estimating the  $P\&L$  that could emerge from an extreme scenario. It is standard to think of that as a 6 standard deviations change in each of the risk factors. For instance, the Profit&Loss effect of an adverse change in the slope of the term structure would be:

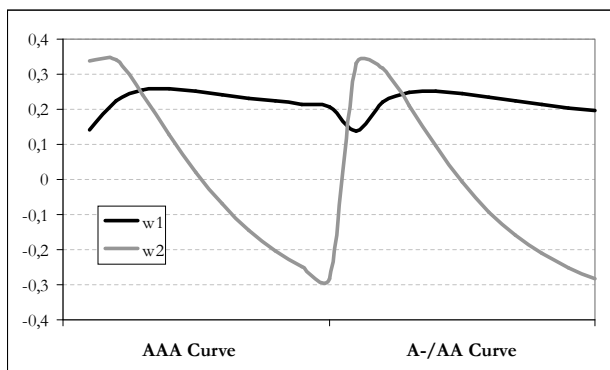
$$Effect_2 = \beta_2 6 \sqrt{12 \lambda_2} = -\text{€}632,415$$

where  $\lambda_2$  denotes the second eigenvalue of the variance covariance matrix of interest rates. The total impact of an adverse movement in the three principal components simultaneously is estimated as a loss of  $-\text{€}1.937.977$  the value of the portfolio.

**Example 23** Example II.2.7: *Multiple curve factor models: PCA on curves with different credit rating. Consider a portfolio invested on fixed income assets with two different rating levels, AAA and A-/AA. Interest rate spreads for the two credit classes at different maturities are provided in the spreadsheet.<sup>15</sup> Compute principal components as risk factors for such portfolio. R : Since both spread curves are very highly correlated, two factors are enough to explain more than 99% of the joint fluctuation in both spread curves. The first factor*

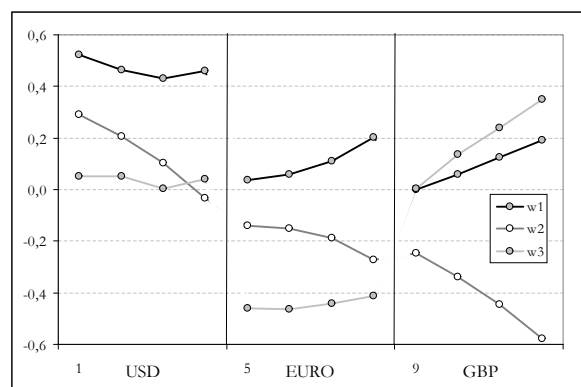
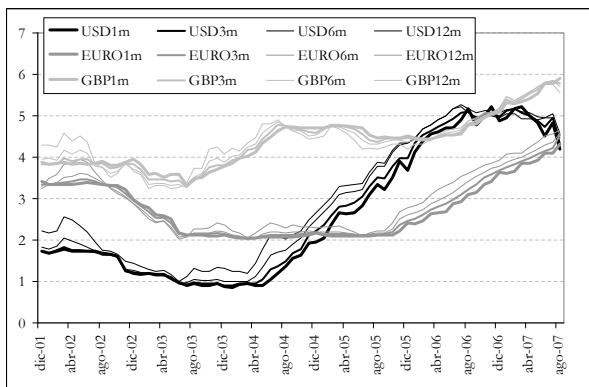
<sup>15</sup>They can be downloaded from <http://sdw.ecb.europa.eu>

represents a parallel shift in both curves. The second factor represents a similar tilt in both curves.



**Example 24** Example II.2.8: Multiple curve factor models: PCA on curves in different currencies

When we examine yield curves in different currencies, it is clear that *within* correlations are much higher than *between* correlations. The spreadsheet contains monthly data on interest rate changes on US Treasury, the euro and the UK yield curves.



The first three eigenvectors are shown in the right Figure. The first eigenvector explains less than 45% of the joint fluctuation, the second eigenvector captures 30% and the third one captures 12% of the fluctuation. We need six components to explain 99% of the joint fluctuation in the three yield curves. The first component is a shift in the US curve, with shifts in the longer end of the euro and UK curves. The second factor is a decrease in short term US rates together with upward moves in euro and UK rates, specially at the longer end. The third factor represents static US when the euro curve shifts down and UK rates tilt up.

## 8.9 Open questions in the analysis of a term structure

### *Cuestiones abiertas*

- Reducing dimensionality of volatility along the term structure of interest rates
- Volatility transmission along the term structure of interest rates
- Term structure of volatilities
- International linkages in the term structure of interest rates
- International transmission of volatility
- Forward rates as predictors of future short-term rates

According to the Expectations Hypothesis, forward rates should be unbiased predictors of future (short-term) spot rates. However, not much effort has been placed to actually test for the predictive ability of forward rates. From the beginning, the statement was tested by fitting regressions like,

$$r_t^m = \beta_0 + \beta_1 f_{t-s,t}^m + u_t$$

and testing the joint hypothesis,

$$H_0 : \beta_0 = 0; \beta_1 = 1$$

but...

- The predictive ability of the term structure

The term structure has been found to contain some useful information regarding future business cycle activity. Initially, it was discovered that the spread between a short and long-term rate could anticipate future output. Such a spread, known as the term structure slope, was included as one of the indicators in the index of leading economic indicators, and has been detected to contain predictive ability in a variety of countries. This empirical result is quite striking, since it is a spread between *nominal* rates which is found to anticipate *real* economic activity.

- General equilibrium characterization of the term structure

Recently, there has been some effort to characterize a term structure of interest rates in stochastic, general equilibrium economies. The idea is to analyze simulated series for interest rates at different maturities, to discuss whether they reproduce some of the regularities observed in actual interest rate data. Some of the difficulties with this exercise are: a) it is hard to solve for such a set of interest rates, since the system easily becomes close to singular, b) to obtain nominal interest rates, we need to use a monetary model, which may be harder to solve.

To this point, most of the work has been devoted to: 1) explain differences in volatility across the term structure [*The term structure of interest rates in real and monetary economies*", Journal of Economic Dynamics and Control 19(5/6), 909-940, W. den Haan (1995)], 2) reproduce ARCH features present in actual interest rate data, 3) reproduce the predictive ability that the term structure seems to contain with respect to the business cycle [*Testing the expectations hypothesis in Eurodeposits*", Journal of International Money and Finance, 19 (2000) 713–736, E.Dominguez y A.Novales (2000)].

## 9 Componentes principales en el análisis de mercados de valores

A similar argument can be applied to solve the asset allocation problem in stock markets. Let us consider a vector time series of international stock market indices, on which a manager would like to diversify his/her position. It is in itself interesting to analyze how much fluctuation in the markets it is common since, at a difference of the TSIR, it is far from obvious that comovements may be very important. Once again, even though the linear combinations that arise as factors may not be easy to interpret, we may get a nice discussion when we succeed in doing so. A specific market index, or the spread between two markets may be some of the principal components arising in this analysis.

Once the components have been characterized, it is clear that a manager investing in them will be capturing the fluctuation in the markets he/she was planning to invest. Notice that each component can be considered as a portfolio. In this case, it will be a portfolio of index portfolios, but the same argument would apply to choosing a portfolio in a given stock market.

An equal weight portfolio might run into important redundancies, thereby producing an inefficient solution to the asset allocation problem. An interesting empirical exercise has to do with following the performance of a portfolio which diversifies among the components, in relation to the performance of the portfolio which is allocated equally among the indices (ver artículos de C.Alexander recomendados para el tema de Cointegración en Econometría Financiera). Not even the diversification among the components is obvious, since we might want to invest more heavily in the first component than in the second, more in this than in the third component, and so on.

A different question relates to the construction of a market index. It is a standard practice that stocks are weighted by capitalization to produce a market index. That implies that in some markets like the Spanish, five or six assets represent 85% of the market index. However, that does not mean that investing in these assets is an efficient strategy, and it might be better to invest in the portfolio defined by the first principal component, or to diversify between the portfolios defined by the first two or three components.

- En su ejemplo 9.1, Tsay calcula CPs de rentabilidades mensuales de IBM, Hewlet-Packard, Intel, Merrill Lynch y Morgan Stanley, de enero 1990 a



diciembre 1999, extendido después hasta diciembre 2003. Utiliza la matriz de correlaciones muestrales, obteniendo que los dos primeros CPs explican el 72% de variabilidad del vector. El primer CP es, aproximadamente, un promedio de los valores, por lo que puede interpretarse como un *componente de mercado*. El segundo CP consta de los tres primeros valores con coeficiente similar y del mismo signo, y de los dos últimos con coeficientes de signo contrario. Por tanto, puede interpretarse como un *componente sectorial* (posición de signo contrario en Tecnología y Finanzas). En este caso, la matriz de covarianzas proporciona resultados similares.

- Con datos de índices de mercados de valores internacionales, es interesante observar cómo los componentes principales pueden recoger promedios de índices por áreas, o diferencias de comportamiento entre ellas: el área USA, el área Latinoamericana, Europa, diferencial Nikkei-Dow Jones, etc.
- En este caso, los componentes principales pueden interpretarse como carteras, proporcionando las ponderaciones que habría que invertir en cada uno de dichos mercados, si el objetivo es estar invertido en todos ellos.
- La ortogonalidad de los componentes elimina la duplicidad en la composición de las carteras, conduciendo así a una gestión eficiente.
- Por razones similares, es interesante preguntarse acerca de la mejor manera de estar invertido en un mercado, como pudiera ser el mercado continuo español. Si con ello queremos decir que intentamos reproducir el índice, es evidente que la solución es compleja; por ejemplo, comprando los 6 valores mayores, de acuerdo con su capitalización, estaríamos generando una correlación muy elevada con el índice. Pero ¿es ésta la manera más eficiente de hacerlo? Es dudoso, por cuanto que un índice generado como media ponderada por capitalización relativa de cada valor, no considera la información proporcionada por las correlaciones entre activos. Esta información es crucial en el cálculo de los componentes principales.
- ¿Cómo debería calcularse un índice de mercado? En función de su finalidad, pero en algunos casos podría considerarse un método de componentes principales, si se pretendiese recoger de un modo eficiente las fluctuaciones en el conjunto de activos que cotizan en dicho mercado.

## 9.1 Equity PCA factor models: [Case Study\_II.2\_PCA Equity factor Model]

Principal components are also very useful to reduce the dimensionality in an analysis of a stock market by computing them to summarize the information on a large vector of returns. They are therefore an essential instrument for risk management and portfolio management in those markets. If we regress the return for each stock on the  $k$  chosen principal components:

$$r_{it} = \alpha_i + \beta_{i1}P_{1t} + \beta_{i2}P_{2t} + \dots + \beta_{ik}P_{kt}, \quad i = 1, 2, \dots, N$$

we can estimate a vector of *betas* and a scalar *alpha* for each stock. Since the Principal Components are obtained from a linear combination of short-term returns, they have  $E(P_{it}) = 0$ , and we then have:

$$\begin{aligned} E(r_{it}) &= \hat{\alpha}_i, \\ \text{Var}(r_{it}) &= \sum_{m=1}^k \beta_{im}^2 \lambda_m, \\ \text{Cov}(r_{it}, r_{jt}) &= \sum_{m=1}^k \beta_{im} \beta_{jm} \lambda_m \end{aligned}$$

where  $\lambda_m$  denotes the  $m$ -th principal component of the covariance/correlation matrix of stock returns, and where we have used the zero correlation property of the principal components. In matrix form, if we denote by  $R$  the vector of  $N$  stock returns, we have:

$$\text{Var}(R) = B\Lambda B'$$

where  $\Lambda$  denotes the  $N \times N$  variance-covariance matrix of the principal components, which will be a diagonal matrix with the eigenvalues  $\lambda$  along the diagonal, and  $B$  is the  $N \times k$  matrix that has in each column the  $k$ -vector of *betas* for each stock.

If we have a portfolio defined by the vector of weights on each stock,  $w' = (w_1, w_2, \dots, w_N)$ , the alpha and beta of the portfolio is obtained from those for the individual stocks by the usual relationships:  $\alpha_c = w'\alpha$  and  $\beta_c = w'B$ . The specific or idiosyncratic risk component of the portfolio is:

$$\text{specific risk} = (w'\Omega w - \beta_c'\Lambda\beta_c)^{1/2}$$

which should be converted to annual volatility using the factor that corresponds to the frequency of the observed data.

Stock data are usually not very highly correlated, so that it is not possible to capture a high percentage of the joint variation with a small number of principal components. However, the number of stocks trading in a given market is huge, so that the need to reduce the dimensionality of the problem is unquestionable.

Principal components become a very useful tool for equity fund management. The spreadsheet contains daily data for 30 stocks trading in the Dow Jones Industrial from 31 December 2004 to 26 April 2006. The first 5 principal components explain 60% of the variance. We also present (pestaña DJIA PC Factor Model) the *representation* model for each stock as a function of the five components, and we can see that the  $R^2$  for the GM stock in those representations is almost 1.

The Principal Components technique allows to easily decompose the total risk in each stock, as well as for any portfolio. We see (pestaña Ex\_II.2.9) that for the equally weighted portfolio (same cash amount invested in each stock) the market risk is equal to the total risk, with an specific risk of zero. For the DJIA index, the systematic risk is 10.05% versus a total risk of 10.11%, with a small specific risk, of only 1.02%. To compute this, we take into account that the DJIA is the portfolio made up of 1 share of each stock, and calculate the weight using prices from 26 April 2006:  $w_{it} = p_{it} / \sum_{i=1}^N p_{jt}$ . We then use these weights, as constant over time, to compute the portfolio betas and the risk decomposition. Notice that this is the right evaluation to make of risk at 26 April 2006 for an investment on DJIA. If we regressed the actual historical time series of the DJIA on the five principal components we would have found a different result, because the DJIA needs rebalancing, as the prices of the different stocks change over time.

## 10 Principal components and cointegration analysis

On the one hand, since interest rates in almost all markets are integrated variables, we have to look at the vector of rates in a large number of maturities as being a vector of (possibly) cointegrated variables. Hopefully, we might find a large number of cointegrating relationships and hence, a small number of common trends among the set of interest rates. If that is the case, then a long-run investor only needs to care about the time behavior of the small set of common trends. Each one of them will be a linear combination of the whole set of interest rates, but still, the exercise reduces to updating the time series of those linear combinations, and follow them.

A difficulty is that to this point, maximum likelihood estimation of the common trends has not yet been included in the standard statistical packages. There is a way of recovering estimated trends using CATS in RATS.

- *Relaciones de cointegración en la Estructura Temporal de Tipos de Interés*

Consideremos el caso de un conjunto de tipos de interés de una estructura temporal corta: tipos de interés a 1, 3, 6 y 12 meses sobre eurodepósitos. En casi todas las divisas, existe evidencia clara acerca de que, siendo los cuatro tipos de interés variables integradas de orden 1,  $I(1)$ , existen 3 relaciones de cointegración entre ellos y, por consiguiente, una única tendencia estocástica común que explica el comportamiento a largo plazo de los cuatro tipos de interés. Es decir, los componentes de largo plazo de los cuatro tipos son esencialmente, el mismo. El primer componente principal del vector de tipos de interés mencionado resulta ser, en todas las divisas, prácticamente un promedio simple de los cuatro tipos de interés; es decir, recoge la evolución temporal del nivel general de la estructura temporal de tipos. Siendo los tipos variables  $I(1)$ , dicho componente es asimismo  $I(1)$ . ¿Son los restantes componentes principales variables estacionarias?

Con estructuras temporales más largas, existen generalmente menos de  $n - 1$  relaciones de cointegración, por lo que es preciso utilizar varios componentes principales. En estos casos es muy relevante la capacidad de los contrastes habituales, incluso multivariantes, de detectar todas las relaciones de cointegración que puedan existir entre las variables.

Es frecuente ver cómo la imposición de restricciones teóricas ayuda a detectar relaciones de cointegración que no se habrían detectado en un análisis estándar. Esta es una cuestión muy relevante en el trabajo empírico.

- *Componentes principales y cointegración: relación entre ambos métodos:*

Dado un vector de variables integradas, algunos de los componentes principales serán asimismo variables integradas. Entre dichas  $M$  variables pueden existir  $r$  relaciones de cointegración y, por tanto,  $M - r$  tendencias diferentes.

Sería interesante que el número de componentes principales no estacionarios coincidiese con el de tendencias distintas. ¿Existe evidencia empírica a este respecto?

De este modo, tendríamos dos representaciones similares, si bien distintas: una, que descompone el (espacio generado por el) vector de variables como combinación lineal de las tendencias y las relaciones de cointegración. Como es conocido, unas y otras, tomadas conjuntamente, generan el mismo espacio que el vector de variables. Si hubiera  $M = m + n$  variables en el vector,  $m$  de las cuales son  $I(1)$ , y  $n$  son  $I(0)$ , existiendo  $r$  relaciones de cointegración entre las variables, sería interesante que  $m - r$  componentes principales fuesen no estacionarios, siendo estacionarios los  $r + n$  componentes principales restantes.

- *Permanent-transitory component decomposition*

(ver Maximum-likelihood decomposition / Granger-Gonzalo decomposition)  
 A short-term investor needs to also worry about the short-term fluctuations in interest rates. So, two issues arise: on the one hand, how to produce an acceptable decomposition of interest rates in permanent-transitory components. Secondly, whether transitory components show high correlations among them. If that is the case, we will still be able to reduce further the dimensionality of the short-run fund manager.

Surprisingly enough, permanent components are able to extract the (stochastic) trend from a set of random variables. That means that, if we have a vector of interest rates, and we are able to characterize a small set of principal components able to explain a large amount of the fluctuations in interest rates, a (linear least-squares) projection on the first principal component is often enough to produce a stationary residual [see *Gourieroux (19xx)*]. That is, each interest rate may be cointegrated with the first principal component. That is not surprising, from how principal components are computed: the first principal component is built so that it captures a high proportion of the variance in the set of variables in the vector. If they are not stationary, and they share a single common trend, the first principal component will not be very different from

that trend. That is the case, in markets for interbank deposits, or euromarkets, where just 1-, 3-, 6- and 12-month rates are usually considered.

In markets where a wider set of maturities is traded, we may need to project on the first two or three principal components to produce stationary residuals, but we might still be able to obtain that trend extraction property. This is a very simple procedure to implement. Its drawback is that part of the stationary component of the series is included in the component which is fitted by the linear combination of principal components estimated by the least-squares projection.

Cuando tan sólo el primer componente principal es no estacionario, una regresión de cada tipo de interés sobre el nivel medio de los mismos, o el primer componente principal, se tiene un residuo estacionario. Esta descomposición entre componentes de largo y corto plazo, o entre *componentes permanente y transitorio* de cada tipo de interés está contaminada por el hecho de que el componente de largo plazo puede incorporar elementos de naturaleza estacionaria que, en una descomposición estructural, querríamos dejar en el denominado componente transitorio, pero puede servir como primera aproximación a este tipo de descomposiciones.

Tiene interés caracterizar las correlaciones entre los componentes transitorios de los tipos de interés así caracterizados. Igualmente, es entonces posible contrastar las posibles relaciones de causalidad que puedan venir sugeridas por alguna teoría explicativa de la formación de la estructura temporal.

En este caso, cada tipo de interés puede escribirse como la suma de dos componentes: 1) la tendencia estocástica común al vector de tipos considerados (multiplicada por una constante), que puede interpretarse como el componente permanente del tipo de interés, 2) un componente estacionario, que puede interpretarse como el componente de corto plazo o componente transitorio de dicho tipo de interés. En algunas ocasiones podríamos observar asimismo un componente de tendencia determinista, representable por un polinomio en la variable tiempo ( $t = 1, 2, 3, \dots$ ), de grado reducido, uno o dos a lo sumo.

- *Relaciones de cointegración bajo la Teoría de las Expectativas*

Por ejemplo, la hipótesis de expectativas de la estructura temporal sugiere que en la comparación entre un tipo de interés a corto y otro tipo a largo (3 meses y 6 meses, por ej.) el tipo de interés a largo plazo es la suma del tipo a corto, y de las expectativas actuales de los valores futuros del tipo a corto, a lo largo del horizonte de inversión cubierto por el tipo a largo. Ello conduce a que: 1) el diferencial largo-corto es el mejor predictor del valor futuro del tipo a corto plazo, 2) este diferencial ha de ser estacionario (ver Domínguez y Novales: "*Testing the Expectations Hypothesis in Eurodeposits*").

En efecto, trabajando con tipos continuos [ $r_t^n = \ln(1 + R_t^n)$ ], tenemos,

$$2r_t^6 = r_t^3 + E_t r_{t+3}^3$$

y suponemos que los agentes tiene expectativas racionales, es decir, que utilizan eficientemente la información de que disponen al predecir tipos de interés futuros,

$$r_{t+3}^3 = E_t r_{t+3}^3 + \varepsilon_{t+3}^3, \text{ con } E_t \varepsilon_{t+3}^3 = 0$$

Por tanto,

$$2r_t^6 = r_t^3 + r_{t+3}^3 - \varepsilon_{t+3}^3$$

es decir,

$$r_{t+3}^3 = 2r_t^6 - r_t^3 + \varepsilon_{t+3}^3 = f_{t,t+3}^3 + \varepsilon_{t+3}^3$$

donde hemos utilizado la definicion del tipo *forward*,  $f_{t,t+3}^3 = 2r_t^6 - r_t^3$ . Esta ecuación sugiere estimar el modelo:

$$r_t^3 = \alpha + \beta f_{t-3,t}^{3,6} + \varepsilon_t^3 \quad (20)$$

en el que tendría interés contrastar las hipótesis:  $H_0 : \beta = 1$  (*HE+ER* en versión débil, con posible prima de riesgo constante),  $H'_0 : \alpha = 0, \beta = 1$  (*HE+ER* en versión estricta),  $H''_0 : \beta = 1$  y  $\varepsilon_t^3 \sim MA(2)$ .

También podríamos escribir,

$$r_{t+3}^3 - r_t^3 = 2(r_t^6 - r_t^3) + \varepsilon_{t+3}^3$$

ecuación que muestra cómo el diferencial entre los tipos a 6 y 3 meses, o pendiente de la *ETTI*, resume la información contenido en la estructura temporal en el instante  $t$  acerca de variaciones futuras en los tipos a corto plazo. Si los tipos de interés continuos son  $I(1)$ , entonces el miembro de la izquierda es estacionario, por ser la primera diferencia en  $r_t^3$ . Pero el error de  $\varepsilon_{t+3}^3$  debe ser estacionario, por lo que también ha de ser estacionario el diferencial  $r_t^6 - r_t^3$ .

Eso significa que el vector de tipos  $(r_t^6, r_t^3)$  está cointegrado, con vector de cointegración  $(1, -1)$ .

Pero este argumento es válido para toda diferencia de tipos de interés, por lo que en un conjunto de  $n$  tipos de interés, debe haber  $n - 1$  relaciones de cointegración independientes. Además, por el mismo argumento que acabamos de hacer, dichas relaciones pueden venir representadas como diferenciales (*spreads*): pueden ser los diferenciales de todos los tipos en el vector respecto del tipo de más corto plazo, los diferenciales de todos los tipos respecto del de más largo plazo, o los diferenciales entre cada dos tipos de interés a vencimientos consecutivos.