Dimension Reduction Techniques

We study *dimension reduction* techniques in these notes focusing in particular on principal components analysis (PCA) and factor models. We will generally follow the notation of Chapter 3 of *Quantitative Risk Management* by *McNeil, Frey and Embrechts* (MFE). This chapter contains detailed discussions of these topics and should be consulted if further details are required.

1 Principal Components Analysis

Let $\mathbf{Y} = (Y_1, \ldots, Y_n)^\top$ denote an n-dimensional random vector with variance-covariance matrix, $\boldsymbol{\Sigma}$. In the context of risk management, we take this vector to represent the (normalized) changes, over some appropriately chosen time horizon, of an n-dimensional vector of risk factors. These risk factors could represent security price returns, returns on futures contracts of varying maturities, or changes in spot interest rates, again of varying maturities. The goal of PCA is to construct linear combinations

$$P_i \;=\; \sum_{j=1}^{n} w_{ij}\, Y_j \quad \text{for} \quad i = 1, \ldots, n$$

in such a way that:

(1) the P_i's are *orthogonal* so that $\mathrm{E}[P_i\, P_j] = 0$ for $i \neq j$ and

(2) the P_i's are ordered so that: (i) P_1 explains the largest percentage of the total variability in the system and (ii) each P_i explains the largest percentage of the total variability in the system that has *not* already been explained by $P_1, \ldots, P_{i-1}$.

In practice it is common to apply PCA to *normalized*[1] random variables that satisfy $\mathrm{E}[Y_i] = 0$ and $\mathrm{Var}(Y_i) = 1$. This is achieved by subtracting the means from the original random variables and dividing by their standard deviations. This is done to ensure that no one component of $\mathbf{Y}$ can influence the analysis by virtue of that component's measurement units. We will therefore assume that the Y_i's have already been normalized. The key tool of PCA is the *spectral decomposition* from linear algebra which states that any symmetric matrix, $\mathbf{A} \in \mathbb{R}^{n\times n}$, can be written as

$$\mathbf{A} \;=\; \boldsymbol{\Gamma}\, \boldsymbol{\Delta}\, \boldsymbol{\Gamma}^\top \tag{1}$$

where:

(i) $\boldsymbol{\Delta}$ is a diagonal matrix, $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, of the *eigen-values* of $\mathbf{A}$ which, without loss of generality, are ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and

(ii) $\boldsymbol{\Gamma}$ is an orthogonal matrix with the i^{th} column of $\boldsymbol{\Gamma}$ containing the i^{th} standardized[2] eigen-vector, γ_i, of $\mathbf{A}$. The orthogonality of $\boldsymbol{\Gamma}$ implies $\boldsymbol{\Gamma}\, \boldsymbol{\Gamma}^\top = \boldsymbol{\Gamma}^\top\, \boldsymbol{\Gamma} = \mathbf{I}_n$.

Since $\boldsymbol{\Sigma}$ is symmetric we can take $\mathbf{A} = \boldsymbol{\Sigma}$ in (1) and the positive semi-definiteness of $\boldsymbol{\Sigma}$ implies $\lambda_i \geq 0$ for all $i = 1, \ldots, n$. The *principal components* of $\mathbf{Y}$ are then given by $\mathbf{P} = (P_1, \ldots, P_n)$ satisfying

$$\mathbf{P} \;=\; \boldsymbol{\Gamma}^\top\, \mathbf{Y}. \tag{2}$$

Note that:

[1] Working with normalized random variables is equivalent to working with the correlation matrix of the un-normalized variables.

[2] By standardized we mean $\gamma_i^\top \gamma_i = 1$.

(a) $\mathrm{E}[P] = \mathbf{0}$ since $\mathrm{E}[Y] = \mathbf{0}$ and

(b) $\mathrm{Cov}(\mathbf{P}) = \mathbf{\Gamma}^\top \mathbf{\Sigma} \mathbf{\Gamma} = \mathbf{\Gamma}^\top (\mathbf{\Gamma} \mathbf{\Delta} \mathbf{\Gamma}^\top) \mathbf{\Gamma} = \mathbf{\Delta}$ so that the components of $\mathbf{P}$ are uncorrelated and

$$\mathrm{Var}(P_i) = \lambda_i. \tag{3}$$

This is consistent with (1) above.

The matrix $\mathbf{\Gamma}^\top$ is called the matrix of factor *loadings*. Note that we can invert (2) to obtain

$$\mathbf{Y} = \mathbf{\Gamma} \mathbf{P}. \tag{4}$$

We can measure the ability of the first few principal components to explain the total variability in the system. We see from (3) that

$$\sum_{i=1}^{n} \mathrm{Var}(P_i) = \sum_{i=1}^{n} \lambda_i = \mathsf{trace}(\Sigma) = \sum_{i=1}^{n} \mathrm{Var}(Y_i) \tag{5}$$

where we have used the fact that the *trace* of a matrix, i.e. the sum of its diagonal elements, is also equal to the sum of its eigen-values. If we take $\sum_{i=1}^{n} \mathrm{Var}(P_i) = \sum_{i=1}^{n} \mathrm{Var}(Y_i)$ to measure the total variability, then by (3) we may interpret the ratio

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i}$$

as measuring the percentage of the total variability that is explained by the first k principal components. This is consistent with (2) above since the λ_i's are non-increasing. In particular, it is possible to show that the first principal component, $P_1 = \gamma_1^\top \mathbf{Y}$, satisfies

$$\mathrm{Var}(\gamma_1^\top \mathbf{Y}) = \max_{\mathbf{a}} \left\{ \mathrm{Var}(\mathbf{a}^\top \mathbf{Y}) \; : \; \mathbf{a}^\top \mathbf{a} = 1 \right\}.$$

Moreover, it is also possible to show that each successive principal component, $P_i = \gamma_i^\top \mathbf{Y}$, satisfies the same optimization problem but with the added constraint that it be orthogonal, i.e. uncorrelated, to $P_1, \ldots, P_{i-1}$.

In financial applications, it is often the case that just two or three principal components are sufficient to explain anywhere from 60% to 95% or more of the total variability. Moreover, it is often possible to interpret the first two or three components. For example, if $\mathbf{Y}$ represents (normalized) changes in the spot interest rate for n different maturities, then the first principal component can usually be interpreted as an (approximate) parallel shift in the spot rate curve, whereas the second component represents a flattening or steepening of the curve. In equity applications, the first component often represents a systematic market factor that impacts all of the stocks whereas the second (and possibly other) components may be identified with industry specific factors.

1.1 Empirical PCA

In practice we do not know the true variance-covariance matrix but it may be estimated using historical data. Suppose then that we have the multivariate observations, $\mathbf{X}_1, \ldots \mathbf{X}_m$, where $\mathbf{X}_t = (X_{t1}, \ldots, X_{tn})^\top$, represents the date t sample observation. It is important that these observations come from a *stationary*[3] time series such as asset returns or yield changes. However $\mathbf{X}_t$ should not represent a vector of price levels, for example, as the latter generally constitute *non-stationary* time series. If μ_j and σ_j for $j = 1, \ldots, n$, are the sample mean and standard deviation, respectively, of $\{X_{tj} \; : \; t = 1, \ldots, m\}$, then we can normalize the data by setting

$$Y_{tj} = \frac{X_{tj} - \mu_j}{\sigma_j} \quad \text{for} \quad t = 1, \ldots, m \quad \text{and} \quad j = 1, \ldots, n.$$

Let $\mathbf{\Sigma}$ denote the sample[4] variance-covariance matrix of the $\mathbf{Y}_t$'s where $\mathbf{Y}_t = (Y_{t1}, \ldots, Y_{tn})^\top$. Then

$$\mathbf{\Sigma} = \frac{1}{m} \sum_{t=1}^{m} \mathbf{Y}_t \, \mathbf{Y}_t^\top$$

[3]If the observations are not drawn from a stationary time-series then it makes no sense to talk about the variance-covariance matrix, $\mathbf{\Sigma}$, of $\mathbf{X}$.

[4]Usually we would write $\widehat{\mathbf{\Sigma}}$ for a sample covariance but we will stick with $\mathbf{\Sigma}$ here.

and we compute the principal components using this covariance matrix. From (4), we see that the original data is obtained from the principal components as

$$\begin{aligned} \mathbf{X}_t &= \text{diag}(\sigma_1, \ldots, \sigma_n)\, \mathbf{Y}_t \;+\; \mu \\ &= \text{diag}(\sigma_1, \ldots, \sigma_n)\, \mathbf{\Gamma}\, \mathbf{P}_t \;+\; \mu \end{aligned} \tag{6}$$

where $\mathbf{P}_t \;:=\; (P_{t1}, \ldots, P_{tn})^\top \;=\; \mathbf{\Gamma}^\top \mathbf{Y}_t$ is the t^{th} sample principal component vector.

1.2 Applications of PCA in Finance

There are many applications of PCA in finance and risk management. They include:

1. **Building Factor Models**

   If we believe the first k principal components explain a sufficiently large amount of the total variability then we may partition the $n \times n$ matrix $\mathbf{\Gamma}$ according to $\mathbf{\Gamma} \;=\; [\mathbf{\Gamma}_1\ \mathbf{\Gamma}_2]$ where $\mathbf{\Gamma}_1$ is $n \times k$ and $\mathbf{\Gamma}_2$ is $n \times (n-k)$. Similarly we can write $\mathbf{P}_t = [\mathbf{P}_t^{(1)}\ \mathbf{P}_t^{(2)}]^\top$ where $\mathbf{P}_t^{(1)}$ is $k \times 1$ and $\mathbf{P}_t^{(2)}$ is $(n-k) \times 1$. We may then use (6) to write

   $$\mathbf{X}_{t+1} \;=\; \mu \;+\; \text{diag}(\sigma_1, \ldots, \sigma_n)\, \mathbf{\Gamma}_1\, \mathbf{P}_{t+1}^{(1)} \;+\; \epsilon_{t+1} \tag{7}$$

   where $\epsilon_{t+1} \;:=\; \text{diag}(\sigma_1, \ldots, \sigma_n)\, \mathbf{\Gamma}_2\, \mathbf{P}_{t+1}^{(2)}$ now represents an *error* term. We can interpret (7) as a k-factor model for the changes in risk factors, $\mathbf{X}_{t+1}$. Note, however, that the components of ϵ_{t+1} are not independent which would be the case in a typical factor model. (See Section 2 below for a very brief introduction to factor models.)

2. **Scenario Generation**

   It is easy to generate scenarios using PCA. Suppose today is date t and we want to generate scenarios over the period $[t, t+1]$. We can then use (7) to apply stresses to the first few principal components, either singly or jointly, to generate loss scenarios. Moreover, we know that $\text{Var}(P_i) = \lambda_i$ and so we can easily control the severity of the stresses.

3. **Estimating VaR and CVaR**

   We can use the model in (7) and Monte-Carlo to simulate portfolio returns. This could be done by simply estimating the joint distribution of the first k principal components. Since they are uncorrelated by construction and we know their variances we could, for example, assume

   $$\mathbf{P}_{t+1}^{(1)} \sim \text{MN}_k(\mathbf{0},\ \text{diag}(\lambda_1, \ldots, \lambda_k))$$

   although we might do much better by assuming a heavy-tailed distribution for $\mathbf{P}_{t+1}^{(1)}$. If we wanted to estimate the *conditional* loss distribution (as is usually the case) then we could use time series methods such as GARCH models to do this.

4. **Portfolio Immunization**

   It is also possible to hedge or immunize a portfolio against moves in the principal components. For example, suppose we wish to hedge the value of a portfolio against movements in the first k principal components. Let V_t be the time t value of the portfolio and assume that our hedge will consist of positions, ϕ_i, in the securities with time t prices, S_{ti}, for $i = 1, \ldots, k$. Let V_t^* denote the time t value of the hedged portfolio so that[5]

$$V_t^* \;=\; V_t \;+\; \sum_{i=1}^{k} \phi_i S_i \tag{8}$$

As usual, let $Z_{(t+1)j}$ denote the date $t+1$ level of the j^{th} risk factor so that $\Delta Z_{(t+1)j} = X_{(t+1)j}$. If the change in value of the *hedged* portfolio between dates t and $t+1$ is denoted by ΔV_{t+1}^*, then we have

$$\begin{aligned}
\Delta V_{t+1}^* \;&\approx\; \sum_{j=1}^{n} \left(\frac{\partial V_t}{\partial Z_{tj}} \;+\; \sum_{i=1}^{k} \phi_i \, \frac{\partial S_{ti}}{\partial Z_{tj}} \right) \, \Delta Z_{(t+1)j} \\
&=\; \sum_{j=1}^{n} \left(\frac{\partial V_t}{\partial Z_{tj}} \;+\; \sum_{i=1}^{k} \phi_i \, \frac{\partial S_{ti}}{\partial Z_{tj}} \right) \, X_{(t+1)j} \qquad (9) \\
&\approx\; \sum_{j=1}^{n} \left(\frac{\partial V_t}{\partial Z_{tj}} \;+\; \sum_{i=1}^{k} \phi_i \, \frac{\partial S_{ti}}{\partial Z_{tj}} \right) \left(\mu_j \;+\; \sigma_j \sum_{l=1}^{k} \mathbf{\Gamma_{jl}} P_l \right) \\
&=\; \sum_{j=1}^{n} \left(\frac{\partial V_t}{\partial Z_{tj}} \;+\; \sum_{i=1}^{k} \phi_i \, \frac{\partial S_{ti}}{\partial Z_{tj}} \right) \, \mu_j \qquad (10) \\
&\quad + \sum_{l=1}^{k} \left(\sum_{j=1}^{n} \left(\frac{\partial V_t}{\partial Z_{tj}} \;+\; \sum_{i=1}^{k} \phi_i \, \frac{\partial S_{ti}}{\partial Z_{tj}} \right) \, \sigma_j \mathbf{\Gamma_{jl}} \right) \, P_l \qquad (11)
\end{aligned}$$

where, ignoring ϵ, we used the factor model representation in (7) in going from (9) to (10). We can now use (11) to hedge the risk associated with the first k principal components: we simply solve for the ϕ_i's so that the coefficients of the P_l's in (11) are zero. This is a system of k linear equations in k unknowns and so it is easily solved.

If we include an additional hedging asset then we could also, for example, ensure that the deterministic component of ΔV_{t+1}^*, i.e. the term in (10), is also zero so that the change in value of the hedged portfolio is (approximately) zero.

Consider a bond portfolio, for example, containing various default-free treasury securities. Then n will be the number of different bond maturities in the portfolio with Z_{tj} denoting the spot-rate for the j^{th} maturity at time t. Then changes in the first three principal components will typically then correspond to a parallel shift, a steepening and a twisting, respectively, in the spot-rate curve. We could take three treasury securities of different maturities as our hedging securities.

Example 1 (Example 18.2 in Ruppert and Matteson)
This example uses daily yields on U.S. Treasury bonds at 11 maturities: T = 1, 3, and 6 months, and 1, 2, 3, 5, 7, 10, 20, and 30 years. In order to analyze daily changes in yields, all 11 time series were differenced and we then applied PCA to study how the yield curves change from day to day. The data was taken from the time period January 2, 1990, to October 31, 2008. For various reasons daily yields were missing from some values of T, however. For example, the 20-year constant maturity series was discontinued at the end of 1986 and reinstated on October 1, 1993. Missing data was handled by simply removing all days with missing values of the differenced data. This left 819 days of data beginning on July 31, 2001 (when the one-month series started) and ending on October 31, 2008, with the exclusion of the period February 19, 2002 to February 2, 2006 when the

[5] If the positions in the hedge securities were funded by borrowing (or possibly lending) in the overnight cash market, say, then we could include this cash position in (8). Note that this cash position will have no exposure to the k factors and so including it will make no difference to how we construct the hedge portfolio, i.e. the ϕ_i's.

30-year Treasury was discontinued. It is clear from Figure 18.1 (a) from Ruppert and Matteson (*SDAFE*, 2015) which displays the yield curve on three dates, that the yield curve can take on a variety of shapes.

The covariance matrix rather than the correlation matrix was used because the variables are comparable and in the same units. The results of the PCA are displayed in Figure 18.1 (b), (c) and (d) from Ruppert and Matteson. The *scree* plot in Figure (b) shows the variances of each of the PC's. It is clear from this that almost all of the total variance is explained using the first 5 PC's. In fact, just the first 3 components explain 94.6% of the total variance. We could reasonably construct a 3-factor model then using just these 3 components and use it to hedge and evaluate the risk of (most) portfolios of U.S. government securities. Moreover, each of the first 3 PC's has an obvious interpretation. In particular, (shocks to) these PC's represent parallel shifts, changes in slope and changes in convexity, respectively, to the yield curve.

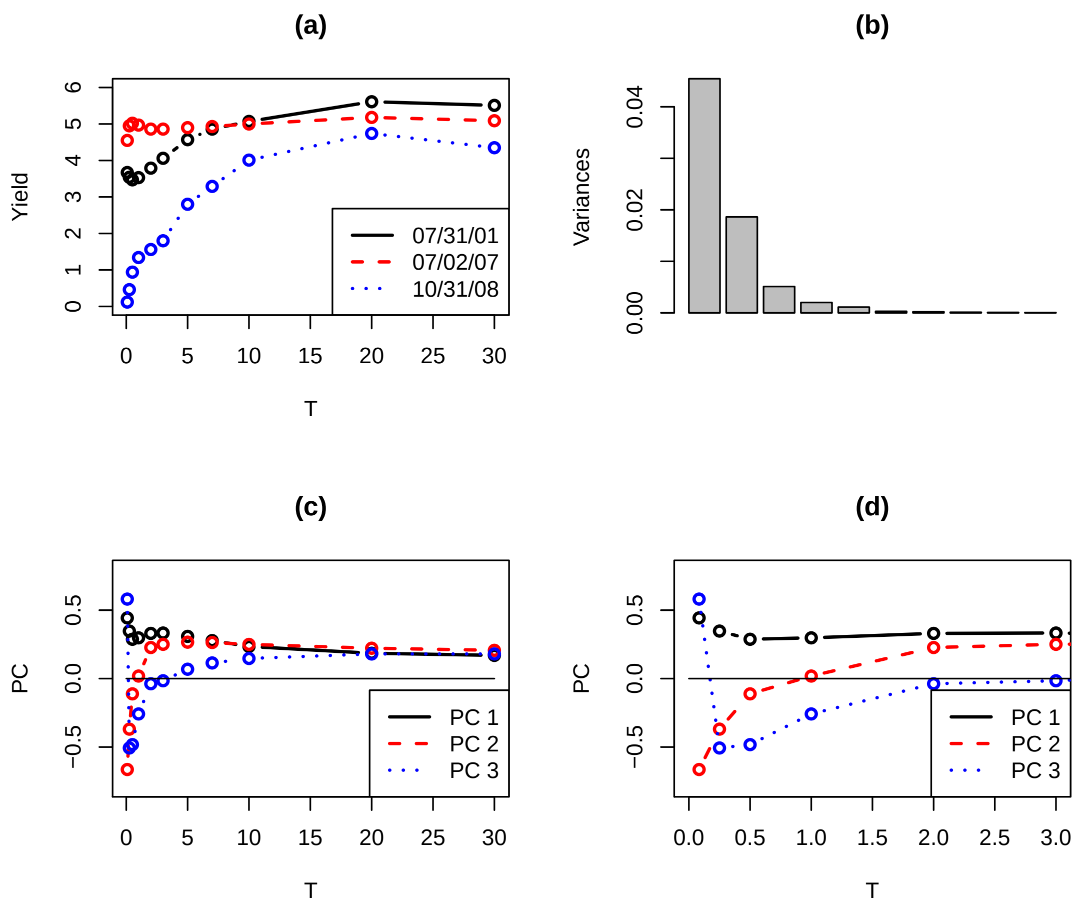

Figure 18.1 from Ruppert and Matteson: (a) Treasury yields on three dates. (b) Scree plot for the changes in Treasury yields. Note that the first three principal components have most of the variation, and the first five have virtually all of it. (c) The first three eigenvectors for changes in the Treasury yields. (d) The first three eigenvectors for changes in the Treasury yields in the range $0 \leq T \leq 3$.

Exercise 1 *Why might some portfolios of U.S. government securities be unsuitable for analysis with such a 3-factor model?*

In Figure 18.2 (a) to (c) of Ruppert and Matteson we can see the mean value of the yield curve plus or minus the first three PC's. These figures emphasize how the yield curve will change with shocks (positive or negative)

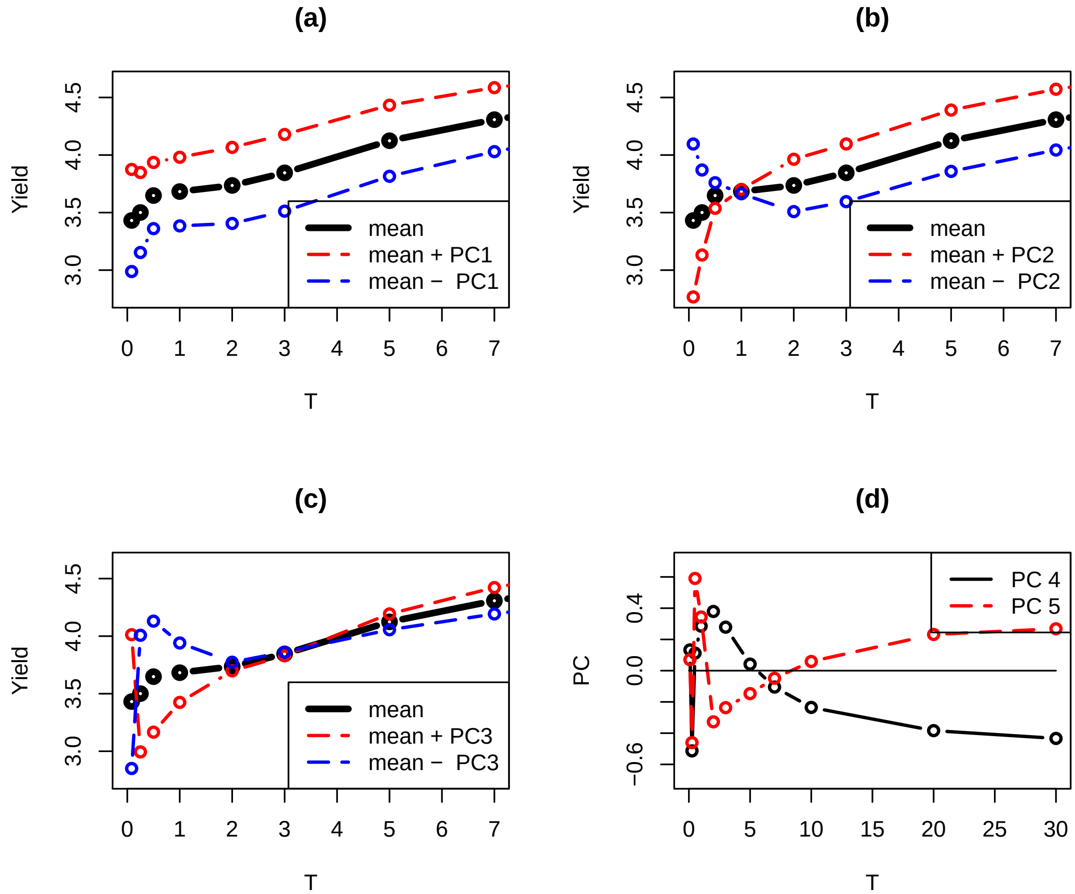

Figure 18.2 from Ruppert and Matteson: (a) The mean yield curve plus and minus the first eigenvector. (b) The mean yield curve plus and minus the second eigenvector. (c) The mean yield curve plus and minus the third eigenvector. (d) The fourth and fifth eigenvectors for changes in the Treasury yields.

to the corresponding PC. But it should be borne in mind that the shocks of ± 1 in the figures are extreme and are only used to aid visualization. In particular, the typical shock size will be determined by the standard deviation of the relevant PC.

In practice we would also be interested in the behavior of the yield changes over time. A time series analysis based on the changes in the 11 yields would be problematic, however, and a better approach would be to use the first three PC's. Their time series and auto- and cross-correlation plots are shown in Figures 18.3 and 18.4 of Ruppert and Matteson, respectively. The auto-correlations and lagged cross-correlations are all quite small and the practical implication of this is that parallel shifts, changes in slopes, and changes in convexity are nearly uncorrelated and could be analyzed separately. Given the volatility clustering that is evident in the time series of Figure 18.3, a reasonable approach would be to model these time series using GARCH models.

Exercise 2 *The lag-0 cross-correlations in Figure 18.4 are zero. Is this just a coincidence? Why or why not?*

∎

It is perhaps worth emphasizing that PCA doesn't always work so well. With many equity portfolios, for example, it is very rare that just 3 PC's will explain so much of the variance. Typically a far greater number are required and, with the exception of the first PC, most of the PC's do not have an obvious interpretation.

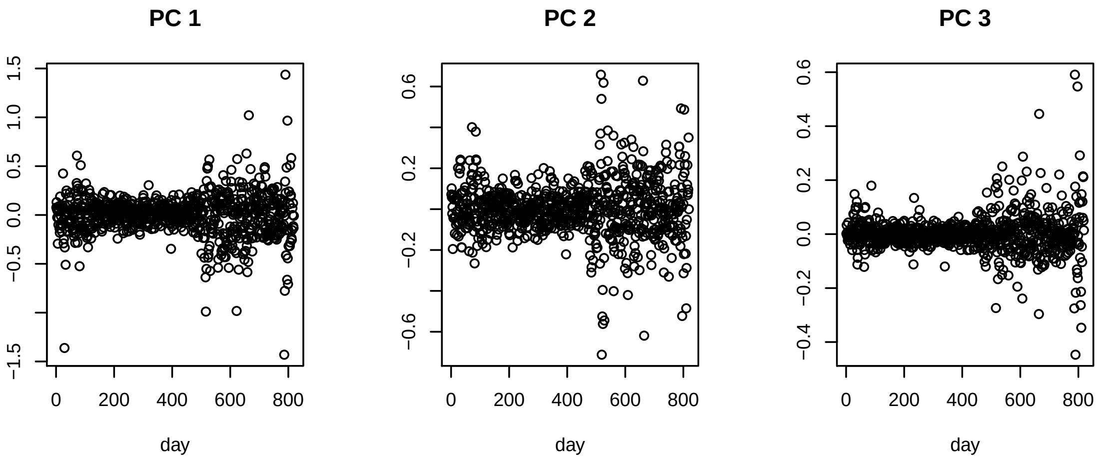

Figure 18.3 from Ruppert and Matteson: Time series plots of the first three principal components of the Treasury yields. There are 819 days of data, but they are not consecutive because of missing data; see text.

Exercise 3 *Can you guess what the first PC typically represents when PCA is applied to equity portfolios?*

---

2 Factor Models

Factor models play an important role in finance, particularly in the equity space where they are often used to build low-dimensional models of stock returns. In this context they are often used for portfolio construction, risk attribution and measurement, portfolio hedging etc. Our discussion is brief[6] and we begin with the definition[7] of a k-factor model:

Definition 1 *We say the random vector* $\mathbf{R}_t := (R_{1,t}, \ldots, R_{n,t})^\top$ *follows a* **linear k-factor model** *if it satisfies*

$$\mathbf{R}_t \;=\; \mathbf{a} \;+\; \mathbf{B}\,\mathbf{F}_t \;+\; \epsilon_t \tag{12}$$

where

(i) $\mathbf{F}_t = (F_{1,t}, \ldots, F_{k,t})^\top$ *is a random vector of* common factors *(or factor returns) with* $k < n$ *and with a positive-definite covariance matrix;*

(ii) $\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_{n,t})$ *is a random vector of idiosyncratic error terms which are uncorrelated and have mean zero;*

(iii) $\mathbf{B}$ *is an* $n \times k$ *constant matrix of* factor loadings, *and* $\mathbf{a}$ *is an* $n \times 1$ *vector of constants;*

(iv) $\mathrm{Cov}(F_{i,t}, \epsilon_{j,t}) = 0$ *for all* i, j.

[6] Chapter 8 of *Market Models* by Carol Alexander (2001), Chapter 17 of Ruppert's *Statistics and Data Analysis for Financial Engineering* or Section 3.4 of *MFE* can be consulted for an introduction to factor models and further references. There are many vendors of factor models for equity markets in the financial services industry. It is debatable as to how much value these models actually provide when it comes to portfolio construction and risk analysis since they generally have relatively little explanatory power.

[7] This is Definition 3.3 in *MFE*.

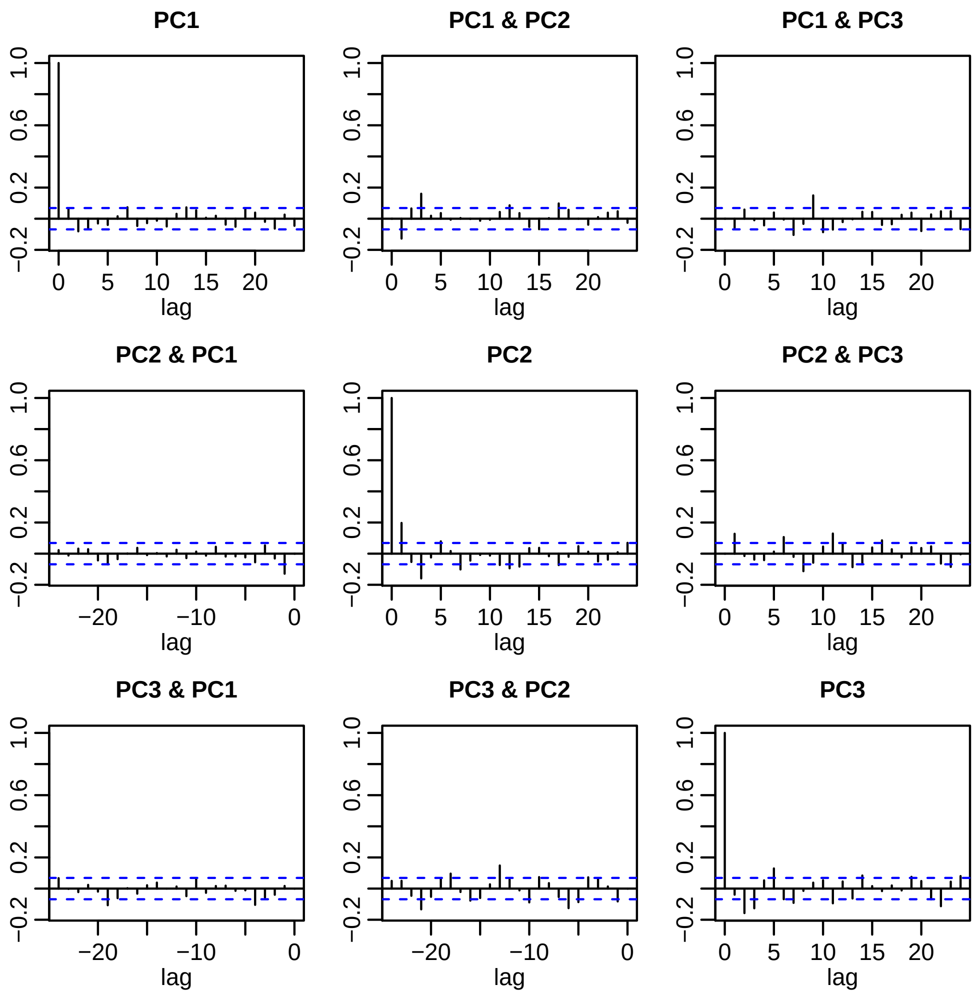

Figure 18.4 from Ruppert and Matteson: Sample auto- and cross-correlations of the first three principal components of the Treasury yields.

Note that t in Definition 1 refers of course to time. We see then that the factors, idiosyncratic error terms and returns, $\mathbf{R}_t$ are stochastic and time-varying whereas all other quantities, i.e. $\mathbf{B}$, μ, $\mathbf{\Sigma}$ etc., are constant. If $\mathbf{R}_t$ is multivariate normally distributed and follows (12) then it is possible to find a version of the model where $\mathbf{F}_t$ and ϵ_t are also multivariate normally distributed. In this case the $\epsilon_{i,t}$'s are independent. If $\mathbf{\Omega}$ is the covariance matrix of $\mathbf{F}_t$ then the covariance matrix, $\mathbf{\Sigma}$, of $\mathbf{R}_t$ must satisfy (why?)

$$\mathbf{\Sigma} \;=\; \mathbf{B}\,\mathbf{\Omega}\,\mathbf{B}^\top \;+\; \mathbf{\Upsilon} \tag{13}$$

where $\mathbf{\Upsilon}$ is a diagonal matrix with $\mathbf{\Upsilon}_{ii} = \mathrm{Var}(\epsilon_{i,t})$.

Exercise 4 *Show that if (12) holds then there is also a representation*

$$\mathbf{R}_t \;=\; \mu \;+\; \mathbf{B}^*\,\mathbf{F}_t^* \;+\; \epsilon_t \tag{14}$$

where $\mathrm{E}[\mathbf{R}_t] = \mu$ *and* $\mathrm{Cov}(\mathbf{F}_t^*) = \mathbf{I}_k$ *so that* $\mathbf{\Sigma} \;=\; \mathbf{B}^*\,(\mathbf{B}^*)^\top \;+\; \mathbf{\Upsilon}$.

Example 2 (Factor Models Based on Principal Components)
The factor model of (7) may be interpreted as a k-factor model with $\mathbf{F}_t = \mathbf{P}_t^{(1)}$ and $\mathbf{B} = \text{diag}(\sigma_1, \ldots, \sigma_n)\, \mathbf{\Gamma}_1$. Note that as constructed, the covariance of the error term, ϵ_t, in (7) is not diagonal and so it does not satisfy part (ii) of our definition above. Nonetheless, it is quite common to construct factor models in this manner and to then make the assumption that ϵ_t is indeed a vector of uncorrelated error terms. ▮

2.1 Taxonomy of Factor Models

According to Ruppert and Matteson (2015) there are three main types of factor models and the calibration approach depends on the type:

1. **Observable Factor Models** are models where the factors, $\mathbf{F}_t$, have been identified in advance and are observable. The factors in these models typically have a fundamental economic interpretation. A classic example would be a 1-factor model where the excess return on the market plays the role of the single factor, $F_{1,t}$. The *Capital Asset Pricing Model* (CAPM) and Fama-French models (see Example 3 below) are examples. In general, potential factors include macro-economic and other financial variables.

   These models are usually calibrated and tested for goodness-of-fit using *multivariate*[8] regression techniques. Alternatively one could simply run n separate regressions using time series data to fit the model, i.e. estimate $\mathbf{B}$ and $\mathbf{\Upsilon}$, in (12). $\mathbf{\Omega}$ (and hence $\mathbf{\Sigma}$) can be estimated directly from the time series observations of $\mathbf{F}_t$. See Section 17.4 of Ruppert for further details and examples.

2. **Cross-Sectional Factor Models** are models where the factors, $\mathbf{F}_t$ are unobserved and the loadings, $\mathbf{B}_t$ are directly observed. This is the exact opposite to the case of an observable factor model as described above. This difference also leads to a different estimation method. Whereas observable factor models can be fit one stock at a time using time series data, a cross-sectional factor model is fit one period at a time using all of the securities. Referring to (1), we can estimate the time t value of $\mathbf{F}_t$ by treating it as a regression problem with n observations corresponding to the returns and loadings for each security at time t. Examples of loadings could be a book-to-market risk factor or a dividend yield factor. In each of these cases the loadings for the j^{th} stock would be the book-to-market value and dividend yield of the j^{th} stock, both of which are directly observable. In this case we would naturally allow $\mathbf{B}$ to vary with time. See Section 17.5 of Ruppert for further details and examples.

3. **Statistical Factor Models** are models where neither the factors values $\mathbf{F}_t$ nor the loadings, $\mathbf{B}$, have been identified in advance. Both therefore need to be estimated as part of the overall statistical analysis. Two standard methods for building such models are *factor analysis* and principal components analysis, the latter of which we have already seen. See Section 17.6 of Ruppert for further details.

Example 3 (The Fama-French 3-Factor Model)
The Fama-French 3-factor model is an observable factor model where the factors[9] are the market return, the excess return (SMB) of small-cap stocks over large-cap stocks, and the excess return (HML) of stocks with a high book-to-price ratio over stocks with a low book-to-price ratio.

The Fama-French 3-factor model assumes that the return $R_{j,t}$ of security j in the j^{th} time period satisfies

$$R_{j,t} = r_{f,t} + \beta_{0,j} + \beta_{1,j}\left(R_{M,t} - r_{f,t}\right) + \beta_{2,j}\mathsf{SMB}_t + \beta_{3,j}\mathsf{HML}_t + \epsilon_{j,t} \tag{15}$$

where $r_{f,t}$, $R_{M,t}$, SMB_t and HML_t are the risk-free rate, market return, SMB and HML factor returns, respectively, in the t^{th} period. Note that these returns are all observable and the j^{th} security loadings on these

[8]A multivariate regression refers to a regression where there is more than one *dependent* variable.

[9]The SMB and HML portfolios change through time and are updated on Ken French's web-page at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research`. Specific details on the factor construction as well as historical factor returns can be found there.

factors, $\beta_{i,j}$ for $i = 0, \ldots, 3$ are unobserved and therefore need to be estimated. As usual, $\epsilon_{j,t}$ is an idiosyncratic error term. Fama and French claim that the most of the pricing anomalies associated with the CAPM 1-factor model disappear in their 3-factor model and this accounts for the popularity of the model in the literature. See Section 18.4.1 in Ruppert and Matteson for further details including an example where they fit the model to real data. ∎

2.2 Factor Models in Risk Management

It is straightforward to use a factor model such as (12) to manage risk. For a given portfolio composition and fixed matrix, $\mathbf{B}$, of factor loadings, the sensitivity of the total portfolio value to each factor, F_i for $i = 1, \ldots, k$, is easily computed. The portfolio composition can then be adjusted if necessary in order to achieve the desired overall factor sensitivity. Obviously this process is easier to understand and justify when the factors are easy to interpret. When this is not the case then the model is purely statistical. This tends to occur when statistical methods such as factor analysis or PCA are used to build the factor model. Of course, as we have seen in the case of PCA, it is still possible even then for the identified factors to have an economic interpretation.

Exercise 5 *Consider a portfolio consisting of λ_i units of the i^{th} stock for $i = 1, \ldots, m$. Suppose the m stocks in the portfolio are drawn from a universe of n stocks with $m \leq n$ and let $\mathbf{R}_t$ represent the vector of $\log$-returns of this universe over some fixed time horizon. If $\mathbf{R}_t$ follows the factor model in (12) determine the sensitivity of the total value of the portfolio to the j^{th} factor, F_j, where $1 \leq j \leq k$.*

Another application of factor models is to the estimation of covariance matrices. In particular, we could estimate $\mathbf{\Sigma}$ using (13) and this approach would have considerable merit when there is only a limited amount of return data available. This is because estimating $\mathbf{\Sigma}$ according to (13) requires the estimation of far fewer parameters than would otherwise be the case if $\mathbf{\Sigma}$ were estimated directly.

In a later set of lecture notes we will also discuss how to *allocate* the risk, e.g. VaR or CVaR, of a portfolio to the various risk factors driving the portfolio return. This will also allow the portfolio / risk manager to obtain a sense of the contributions of each factor to the overall portfolio risk. We finish in the next subsection with a simple yet commonly used approach to scenario analysis based on just a couple of risk factors.

Example 4 (Scenario Analysis Using a Simple Ad-Hoc 2-Factor Model)
John Smith has a portfolio consisting of various stock positions as well as a number of equity options. The portfolio can be seen in the figure below. Mr. Smith would like to perform a basic scenario analysis with just two factors:

1. An equity factor, F_{eq}, representing the equity market.
2. A volatility factor, F_{vol}, representing some general implied volatility factor.

We can perform the scenario analysis by stressing combinations of the factors and computing the P&L resulting from each scenario. Of course using just two factors for such a portfolio will result in a scenario analysis that is quite coarse but in many circumstances this may be sufficient. The question that now arises is how do we compute the value of each security in each scenario? This is easy if we adopt a factor model framework. Consider `AAPL`, for example, and suppose we have estimated β_{eq}^{AAPL} and $\beta_{vol}^{\text{AAPL}}$, the sensitivity of `AAPL`'s stock return and "implied volatility" to the equity and volatility factors, respectively. If ΔF_{eq} is the shock, e.g. $+10\%$ or -5%, to the equity factor then we simply assume that `AAPL` stock experiences a shock of $\beta_{eq}^{\text{AAPL}} \times \Delta F_{eq}$. Similarly the change in implied volatility of `AAPL` will be $\beta_{vol}^{\text{AAPL}} \times \Delta F_{vol}$ where ΔF_{vol} is the shock to the volatility factor. We can then use the Black-Scholes formula to re-value a European option on `AAPL` in that scenario. (An American option can be re-valued numerically using a binomial model for example.) We can therefore compute the P&L on the entire portfolio in each stress scenario of $(\Delta F_{eq}, \Delta F_{vol})$ and construct a table of P&L's as in Figure 1 below.

How do we determine β_{eq}^{AAPL} and $\beta_{vol}^{\text{AAPL}}$? We could use standard statistical methods but we could also simply choose values that we believe are sensible. As is always the case with scenario analysis, performing it looks like a

John Smith's Portfolio

Security Type	Underlying	Underlying Price	Call (C) / Put (P)	Eur / Amer	Div-Yield	Risk-free rate	Maturity	Strike	Implied Vol	Quantity
Stock	MSFT	50.00			0.13%	0.50%				40
Stock	IBM	75.00			0.00%	0.50%				50
Stock	C	10.00			0.24%	0.50%				75
Option	C	10.00	P	Amer	0.24%	0.50%	0.5	10	85%	-500
Option	SP500	150.00	C	Eur	0.00%	0.50%	0.25	150	65%	400
Futures	SP500	150.00			0.00%	0.50%				400
Stock	APPL	100.00			0.26%	0.50%				50
Stock	SP500	150.00			0.00%	0.50%				-35
Cash	EU	1.00			0.00%	0.30%				8000
Option	APPL	100.00	C	Eur	0.26%	0.50%	0.25	110	15%	-200
Option	WMT	40.00	P	Eur	0.26%	0.50%	0.5	25	30%	500
Option	IBM	75.00	P	Amer	0.00%	0.50%	0.25	75	20%	200
Stock	WMT	40.00			0.09%	0.50%				90
Stock	Siemens	35.00			0.11%	0.30%				-100
Option	Siemens	35.00	P	Eur	0.11%	0.30%	1	35	36%	100
Option	APPL	100.00	C	Eur	0.26%	0.30%	0.5	100	40%	-1000
Option	APPL	100.00	P	Eur	0.26%	0.30%	0.5	100	40%	-1000

very straightforward task but choosing sensible risk-factors and stresses does require some work. Note also that in this case the reported P&L's will only be approximations in the sense that even if the factor modeling was correct, the scenario analysis ignores idiosyncratic noise and also makes other implicit assumptions such as assuming that FX rates do not change.

Exercise 6 *Which (if any) of the three classes of factor models described in Section 2.1 best describes this 2-factor model?*

Finally, note that the same factor modeling approach can also be used to construct aggregate Greeks for the portfolio. As before these Greeks should be consistent with the reported P&L's in the scenario table for small or reasonably-sized stresses. ∎

P&L									
0.00	-10%	-5%	-2%	-1%	0%	1%	2%	5%	10%
20%	(778)	(2,900)	(4,295)	(4,773)	(5,256)	(5,744)	(6,236)	(7,727)	(10,250)
-10%	3,006	188	(1,514)	(2,079)	(2,645)	(3,210)	(3,776)	(5,461)	(8,262)
-5%	4,572	1,598	(174)	(764)	(1,354)	(1,943)	(2,533)	(4,287)	(7,192)
-2%	5,473	2,461	650	49	(551)	(1,152)	(1,752)	(3,545)	(6,519)
-1%	5,781	2,757	939	333	(273)	(878)	(1,483)	(3,293)	(6,289)
0%	6,096	3,053	1,221	610	0	(610)	(1,219)	(3,043)	(6,062)
1%	6,417	3,356	1,512	898	283	(331)	(945)	(2,782)	(5,823)
2%	6,743	3,658	1,796	1,178	559	(59)	(677)	(2,525)	(5,587)
5%	7,737	4,576	2,679	2,045	1,411	777	143	(1,744)	(4,867)
10%	9,220	6,040	4,100	3,454	2,807	2,159	1,509	(427)	(3,644)
20%	10,710	8,102	6,353	5,751	5,140	4,524	3,900	2,002	(1,219)

Figure 1: Scenario analysis for John Smith's portfolio. Equity factor stressed from -20% to + 20% and volatility factor stressed from -10% to + 10% volatility points.