CHAPTER 10

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a technique for determining if a high dimensional joint distribution lies approximately in a lower dimensional subspace. Reducing the dimensionality of an application can be useful, especially in a simulation setting.

§1. Basics of PCA

Suppose the $n \times p$ matrix $\mathbf{X} = (\mathbf{X}_{tj})$ is a time series (hence t instead of the customary i) of row vectors of data. Each row vector contains p items of data which are observed over n time periods. Assume that the entries of each column sum to zero, so, of course, the mean of each column is also zero.

Our Application. Following a similar example in [21] (page 44), we will consider daily historical Treasury yield curve data from the period February 2006 to April 2018. The data will comprise the 1, 2, 3, 5, 7, 10, 20, and 30 year Treasury par yields — 3041 observations of the yield curve. These eight numbers reported daily may be viewed as points in $\mathbb{R}^8$ — eight dimensional Euclidean space. Here $n = 3041$ and $p = 8$ and each row of the matrix $\mathbf{X}$ corresponds to the yield curve on one day. Let $\mu_1, \ldots, \mu_p$ denote the "mean yield curve" over this time, so $\mu_j = \frac{1}{n}\sum_{t=1}^{n} \mathbf{X}_{tj}$, $1 \leq j \leq p$. We subtract these means from each column of the data, so the new $\mathbf{X}_{tj}$ is the old $\mathbf{X}_{tj}$ minus μ_j. In this fashion each column is now mean-zero and each row measures how the corresponding yield curve differs from the mean yield curve. We will see that these points all lie quite close to a three dimensional subspace of $\mathbb{R}^8$.

How it Works. Using that the sample mean of each column is zero, we compute $\mathbf{X}$'s $p \times p$ sample covariance matrix $\mathbf{V} = (\mathbf{V}_{ij})$, also denoted $\operatorname{Cov}(\mathbf{X})$, as

$$\mathbf{V}_{ij} = \frac{1}{n}\sum_{t=1}^{n} \mathbf{X}_{ti}\mathbf{X}_{tj} = \frac{1}{n}\sum_{t=1}^{n} \mathbf{X}^T_{it}\mathbf{X}_{tj} = \frac{1}{n}(\mathbf{X}^T\mathbf{X})_{ij},$$

so $\mathbf{V} = \frac{1}{n}\mathbf{X}^T\mathbf{X}$. Let $\lambda_1 > \lambda_2 > \cdots > \lambda_p$ denote $\mathbf{V}$'s eigenvalues listed in descending order and $\mathbf{q}_1, \ldots, \mathbf{q}_p$ denote the corresponding eigenvectors, chosen to be of unit length. Since $\mathbf{V}$ is real, symmetric, and positive definite we know that the eigenvalues are positive and the eigenvectors are orthogonal (See Chapter

3). Put the p normalized eigenvectors into the columns of a matrix $\mathbf{Q}$, so $\mathbf{q}_j =$ column j of $\mathbf{Q}$.

Remark. We actually only know that $\mathbf{V}$ is PSD, but since n is so much larger than p here the likelihood of $\mathbf{V}$ not being PD is slim. For this to happen, we would need at least one column to be a linear combination of the others.

Since the eigenvectors are orthonormal we have that

$$(\mathbf{Q}^T\mathbf{Q})_{ij} \;=\; \mathbf{q}_i^T\mathbf{q}_j \;=\; \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

so $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$, yielding that $\mathbf{Q}^T = \mathbf{Q}^{-1}$ and also $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$. Here $\mathbf{I}$ is the $p \times p$ identity matrix.

Now let $\mathbf{\Lambda}$ be the $p \times p$ matrix with the eigenvalues on the diagonal (in decreasing order) and zeros off the diagonal. Then we have:

Fact 1. $\mathbf{VQ} = \mathbf{Q\Lambda}$, so also $\mathbf{Q}^T\mathbf{VQ} = \mathbf{\Lambda}$.

Proof. The j^{th} column of $\mathbf{VQ}$ is $\mathbf{V}$ times the j^{th} column of $\mathbf{Q}$, i.e., the j^{th} column of $\mathbf{VQ}$ is $\mathbf{V}\mathbf{q}_j$, which is $\lambda_j\mathbf{q}_j$. On the other hand, the j^{th} column of $\mathbf{Q\Lambda}$ is $\mathbf{Q}$ times the j^{th} column of $\mathbf{\Lambda}$. Now the j^{th} column of $\mathbf{\Lambda}$ is a column vector with zeros everywhere except for entry j, which is λ_j. Hence the j^{th} column of $\mathbf{Q\Lambda}$ is λ_j times the j^{th} column of $\mathbf{Q}$, i.e., it is also $\lambda_j\mathbf{q}_j$. Hence $\mathbf{VQ} = \mathbf{Q\Lambda}$. Since $\mathbf{Q}^T = \mathbf{Q}^{-1}$ we get that $\mathbf{Q}^T\mathbf{VQ} = \mathbf{\Lambda}$. $\square$

A Simple Example. Suppose, for example, that $\mathbf{V} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, where $0 < \rho < 1$. We know that $\lambda_1 = 1+\rho$ and $\lambda_2 = 1-\rho$ with normalized eigenvectors $\mathbf{q}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{q}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Here $\mathbf{Q}^T = \mathbf{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and

$$\mathbf{Q}\mathbf{Q}^T \;=\; \mathbf{Q}^T\mathbf{Q} \;=\; \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;=\; \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \;=\; \mathbf{I}. \quad \checkmark$$

Also

$$\begin{aligned} \mathbf{Q}^T\mathbf{VQ} \;&=\; \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &=\; \frac{1}{2}\begin{pmatrix} 1+\rho & 1+\rho \\ 1-\rho & -1+\rho \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &=\; \frac{1}{2}\begin{pmatrix} 2+2\rho & 0 \\ 0 & 2-2\rho \end{pmatrix} \;=\; \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \;=\; \mathbf{\Lambda}. \quad \checkmark \end{aligned}$$

Returning to the general situation, let $\mathbf{Y} = \mathbf{XQ}$ so also $\mathbf{Y}\mathbf{Q}^T = \mathbf{X}$. Then $\mathbf{Y} = (\mathbf{Y}_{tj})$ is $n \times p$ and we verify that $\mathbf{Y}$'s columns are mean 0: For $1 \leq j \leq p$,

$\mathbf{Y}_{tj} = (\mathbf{XQ})_{tj} = \sum_{k=1}^{p} \mathbf{X}_{tk}\mathbf{Q}_{kj}$, so

$$\sum_{t=1}^{n} \mathbf{Y}_{tj} \;=\; \sum_t \sum_{k=1}^{p} \mathbf{X}_{tk}\mathbf{Q}_{kj} \;=\; \sum_k \mathbf{Q}_{kj} \sum_t \mathbf{X}_{tk} \;=\; 0,$$

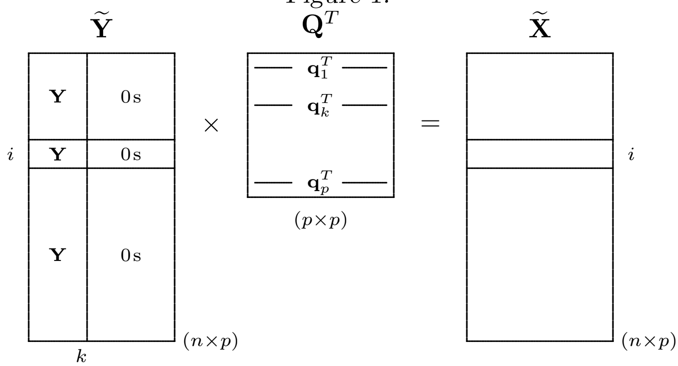
as each $\sum_t \mathbf{X}_{tk}$ is 0. We therefore have

$$\begin{aligned} \mathbf{V} \;=\; \mathrm{Cov}\,(\mathbf{X}) \;&=\; \frac{1}{n}\mathbf{X}^T\mathbf{X} \\ &=\; \frac{1}{n}(\mathbf{YQ}^T)^T(\mathbf{YQ}^T) \;=\; \frac{1}{n}\mathbf{Q}(\mathbf{Y}^T\mathbf{Y})\mathbf{Q}^T \;=\; \mathbf{Q}\mathrm{Cov}\,(\mathbf{Y})\mathbf{Q}^T, \end{aligned}$$

with the final equality justified by the fact that the columns of $\mathbf{Y}$ are mean-zero. Pre- and post-multiplying both sides of this by $\mathbf{Q}^T$ and $\mathbf{Q}$, respectively, gives $\mathrm{Cov}\,(\mathbf{Y}) = \mathbf{Q}^T\mathbf{VQ} = \mathbf{\Lambda}$.

The eigenvalues λ_i are presented in decreasing order. Suppose that, for some number $k < p$, $\sum_{i=1}^{k} \lambda_i$ is a "large portion" of $\sum_{i=1}^{p} \lambda_i$. Let $\widetilde{\mathbf{Y}}$ agree with $\mathbf{Y}$ in columns 1 through k, but with columns $k+1$ through p all zeros and put $\widetilde{\mathbf{X}} = \widetilde{\mathbf{Y}}\mathbf{Q}^T$. Row i in $\widetilde{\mathbf{X}}$, which is a point in $\mathbb{R}^p$, is then a linear combination of the orthogonal eigenvectors $\mathbf{q}_1, \ldots, \mathbf{q}_k$ with weights given by the first k components of the i^{th} row of $\widetilde{\mathbf{Y}}$. (See Figure 1). The n rows of $\widetilde{\mathbf{X}}$ therefore lie in a k-dimensional subspace of $\mathbb{R}^p$. In particular, the i^{th} row of $\widetilde{\mathbf{X}}$ is the orthogonal projection of the i^{th} row of $\mathbf{X}$ onto the vector space $\mathrm{Span}\,\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$. It is easily verified that the columns of $\widetilde{\mathbf{X}}$ are mean-zero.

Figure 1.



Presently we verify that the rows in $\widetilde{\mathbf{X}}$ have almost the same covariance structure as the rows in $\mathbf{X}$. To see this, the reader will verify that

$$\mathrm{Cov}\,(\widetilde{\mathbf{Y}}) \;=\; \frac{1}{n}\widetilde{\mathbf{Y}}^T\widetilde{\mathbf{Y}} \;=\; \widetilde{\mathbf{\Lambda}},$$

where $\widetilde{\mathbf{\Lambda}}_{ii} = \lambda_i$ for $1 \le i \le k$ with zeros elsewhere. Then

$$\begin{aligned}
\operatorname{Cov}(\mathbf{X}) - \operatorname{Cov}(\widetilde{\mathbf{X}}) &= \frac{1}{n}(\mathbf{X}^T\mathbf{X} - \widetilde{\mathbf{X}}^T\widetilde{\mathbf{X}}) \\
&= \frac{1}{n}\mathbf{Q}(\mathbf{Y}^T\mathbf{Y} - \widetilde{\mathbf{Y}}^T\widetilde{\mathbf{Y}})\mathbf{Q}^T \\
&= \mathbf{Q}(\mathbf{\Lambda} - \widetilde{\mathbf{\Lambda}})\mathbf{Q}^T \ = \ \mathbf{Q}\mathbf{D}\mathbf{Q}^T,
\end{aligned}$$

where the diagonal matrix $\mathbf{D}$ has $\mathbf{D}_{ii} = \lambda_i$ for $k < i \le p$ with zeros elsewhere. Letting $Q_{\max} = \max_{ij} |\mathbf{Q}_{ij}|$, we have

$$\begin{aligned}
\left| \left[\operatorname{Cov}(\mathbf{X}) - \operatorname{Cov}(\widetilde{\mathbf{X}})\right]_{ij} \right| &= \left| \sum_{a=1}^{p}\sum_{b=1}^{p} \mathbf{Q}_{ia}\mathbf{D}_{ab}\mathbf{Q}^T_{bj} \right| \\
&= \left| \sum_{a} \mathbf{Q}_{ia}\mathbf{D}_{aa}\mathbf{Q}_{ja} \right| \\
&\le Q^2_{\max} \sum_{a} |\mathbf{D}_{aa}| \ = \ Q^2_{\max} \cdot (\lambda_{k+1} + \cdots + \lambda_p).
\end{aligned}$$

The second equality holds since the only term in the inner summation that is non-zero is when $b = a$ as $\mathbf{D}$ is diagonal. Recall that the columns of $\mathbf{Q}$ are unit vectors implying that $Q_{\max} \le 1$. Under our assumption that $\lambda_1 + \cdots + \lambda_k$ is a large portion of the total sum $\lambda_1 + \cdots + \lambda_p$, the quantity on the right (and hence left) will be small.

§2. Back to the Treasury Example

Recall that our data $\mathbf{X}$ consists of 3041 daily observations of the Treasury yield curve (comprising the 1, 2, 3, 5, 7, 10, 20, and 30 year Treasury yields) with the mean of each column subtracted off so that the columns are mean-zero. These eight means are 1 year 1.22%, 2 year 1.43%, 3 year 1.66%, 5 year 2.15%, 7 year 2.56%, 10 year 2.93%, 20 year 3.49%, and 30 year 3.67%. The sample standard deviations about these means are 1 year 1.62%, 2 year 1.48%, 3 year 1.38%, 5 year 1.22%, 7 year 1.10%, 10 year 1.01%, 20 year 0.97%, and 30 year 0.83%. The sample correlations of the yields for the eight maturities are given by:

$$\mathbf{R} = \begin{pmatrix}
1.00 & 0.99 & 0.98 & 0.94 & 0.88 & 0.81 & 0.68 & 0.64 \\
0.99 & 1.00 & 0.99 & 0.96 & 0.91 & 0.85 & 0.71 & 0.67 \\
0.98 & 0.99 & 1.00 & 0.98 & 0.94 & 0.88 & 0.75 & 0.71 \\
0.94 & 0.96 & 0.98 & 1.00 & 0.99 & 0.95 & 0.84 & 0.80 \\
0.88 & 0.91 & 0.94 & 0.99 & 1.00 & 0.98 & 0.91 & 0.88 \\
0.81 & 0.85 & 0.88 & 0.95 & 0.98 & 1.00 & 0.97 & 0.95 \\
0.68 & 0.71 & 0.75 & 0.84 & 0.91 & 0.97 & 1.00 & 0.99 \\
0.64 & 0.67 & 0.71 & 0.80 & 0.88 & 0.95 & 0.99 & 1.00
\end{pmatrix}.$$

A Few Observations. On average, the yield curve was positively sloped over this time period, rising monotonically from 1.22% for the one year yield to 3.67% for the 30 year yield. The variability of the yields decreases monotonically from the short end to the long end. Correlations in $\mathbf{R}$ decrease as you move away from the diagonal, so, for example, the 1 and 30 year yields are the least correlated at 0.64. All of these qualitative features are typical and would be observed over any time period of sufficient length.

In this example, the three largest eigenvalues for the covariance matrix $\mathbf{V}$ (not shown above) are $\lambda_1 = 11.04$, $\lambda_2 = 0.97$, and $\lambda_3 = 0.08$. The sum of the remaining five eigenvalues is 0.01, so these three eigenvalues do account for a "large portion" of $\sum_{i=1}^{p} \lambda_i$. The corresponding eigenvectors are:

$$\begin{aligned}
\mathbf{q}_1 &= (+0.47,\ +0.44,\ +0.41,\ +0.36,\ +0.32,\ +0.28,\ +0.25,\ +0.20)^T, \\
\mathbf{q}_2 &= (-0.42,\ -0.31,\ -0.19,\ +0.05,\ +0.21,\ +0.35,\ +0.53,\ +0.49)^T, \text{ and} \\
\mathbf{q}_3 &= (+0.54,\ +0.08,\ -0.26,\ -0.47,\ -0.42,\ -0.07,\ +0.29,\ +0.37)^T.
\end{aligned}$$

As reported above, the vector of means $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_8)^T$ is:

$$\boldsymbol{\mu} = (1.22,\ 1.43,\ 1.66,\ 2.15,\ 2.56,\ 2.93,\ 3.49,\ 3.67)^T.$$

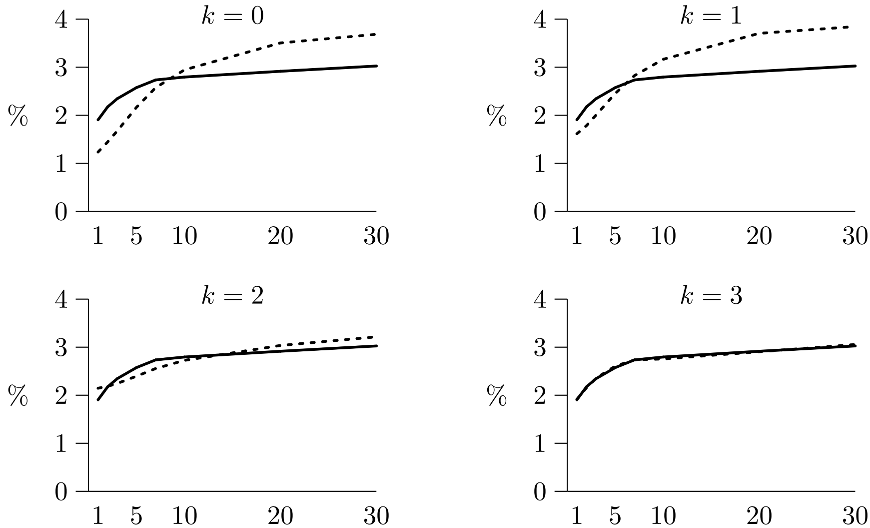
The three eigenvectors above explain almost all the variation in shape of the Treasury yield curve about the mean $\boldsymbol{\mu}$. The vector $\mathbf{q}_1$ corresponds to a *shift* in the yield curve as all the components have the same sign (not quite a "parallel" shift). The vector $\mathbf{q}_2$ corresponds to a *twist* of the curve, as the components go mostly monotonically from negative to positive. The vector $\mathbf{q}_3$ corresponds to a changing yield curve *hump* for obvious reasons. (Perhaps not so obvious — $\mathbf{q}_3$ looks more like a valley than a hump. But if its coefficient is negative, as it is for the example below, then it indeed represents a hump.) Some refer to this as the *curvature* of the yield curve.

The figures below illustrate this for the Treasury yield curve on February 1, 2018 (this is row 3000 of the data in $\mathbf{X}$). For $k = 0$, none of the eigenvectors enter in and the actual curve (solid) and the mean curve $\boldsymbol{\mu}$ (dashed) are shown (with linear interpolation between the eight points). Now row 3000 of the $\mathbf{Y}$ matrix is

$$(\,0.79 \quad -1.27 \quad -0.43 \quad 0.03 \quad -0.02 \quad 0.05 \quad 0.03 \quad 0.01\,),$$

and, when $k = 1$, the dashed curve shows $\boldsymbol{\mu} + 0.79\mathbf{q}_1$. This is $\boldsymbol{\mu}$ shifted up slightly. When $k = 2$, the dashed curve shows $\boldsymbol{\mu} + 0.79\mathbf{q}_1 - 1.27\mathbf{q}_2$. This is $\boldsymbol{\mu}$ shifted up and twisted clockwise. When $k = 3$, the dashed curve shows $\boldsymbol{\mu} + 0.79\mathbf{q}_1 - 1.27\mathbf{q}_2 - 0.43\mathbf{q}_3$. This is $\boldsymbol{\mu}$ shifted up, twisted clockwise, with added hump (as $\mathbf{q}_3$'s coefficient, -0.43, is negative). Here the approximation is almost perfect. The numbers 0.79, -1.27, and -0.43 are also the first three entries in row 3000 of the $\widetilde{\mathbf{Y}}$ matrix with $k = 3$.

Figure 2. The Treasury example with $k = 0$, 1, 2, and 3.



Another Interpretation. A seemingly different calculation yields the same results. Specifically, consider the ordinary least square model

$$\begin{bmatrix} 1.22 \\ 1.43 \\ 1.66 \\ 2.15 \\ 2.56 \\ 2.93 \\ 3.49 \\ 3.67 \end{bmatrix} + \begin{bmatrix} 0.47 & -0.42 & 0.54 \\ 0.44 & -0.31 & 0.08 \\ 0.41 & -0.19 & -0.26 \\ 0.36 & 0.05 & -0.47 \\ 0.32 & 0.21 & -0.42 \\ 0.28 & 0.35 & -0.07 \\ 0.25 & 0.53 & 0.29 \\ 0.20 & 0.49 & 0.37 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1.89 \\ 2.16 \\ 2.33 \\ 2.56 \\ 2.72 \\ 2.78 \\ 2.90 \\ 3.01 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \end{bmatrix}, \qquad (1)$$

where the column vector of numbers on the right is the Treasury yield curve on February 1, 2018 and the β's are selected to minimize $\sum_{i=1}^{8} \epsilon_i^2$. Writing (1) more compactly as $\boldsymbol{\mu} + \mathbf{A}\boldsymbol{\beta} = \mathbf{y} + \boldsymbol{\epsilon}$, we recall that the OLS solution to this is

$$\boldsymbol{\beta} \;=\; (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T(\mathbf{y} - \boldsymbol{\mu}).$$

Since the eigenvectors are orthonormal we see that $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ — the 3×3 identity matrix — so $\boldsymbol{\beta} = \mathbf{A}^T(\mathbf{y} - \boldsymbol{\mu})$, or $\boldsymbol{\beta}^T = (\mathbf{y} - \boldsymbol{\mu})^T\mathbf{A}$. Now $(\mathbf{y} - \boldsymbol{\mu})^T$ is row 3000 of the matrix $\mathbf{X}$ while the matrix $\mathbf{A}$ holds the first three columns of $\mathbf{Q}$. Recalling that $\mathbf{Y} = \mathbf{XQ}$, it follows that the three entries of the row vector $\boldsymbol{\beta}^T$ are the first three entries of row 3000 of the matrix $\mathbf{Y}$. That is, $\beta_1 = 0.79$, $\beta_2 = -1.27$, and $\beta_3 = -0.43$.

It should not be surprising that these two calculations yield the same result. In both cases $\beta_1\mathbf{q}_1 + \beta_2\mathbf{q}_2 + \beta_3\mathbf{q}_3$ is the orthogonal projection of the vector $\mathbf{y} - \boldsymbol{\mu}$ onto the subspace spanned by the vectors $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$.

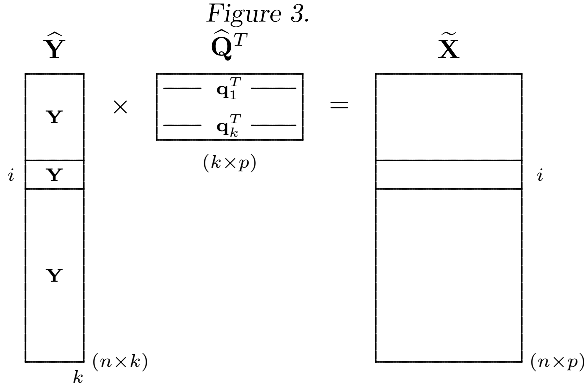
§3. PCA and Simulation — The Treasury Example

Suppose we wish to simulate Treasury yield curves drawn from a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{V}$ as calculated from the data in $\mathbf{X}$ from our Treasury example. Without using PCA, we can do this by calculating the Cholesky decomposition $\mathbf{V} = \mathbf{L}\mathbf{L}^T$. For each simulation we then generate a vector $\mathbf{N} = (N_1, \ldots, N_8)^T$ of independent standard normals, so $\mathbf{LN}$ is an eight dimensional multivariate normal with mean $\mathbf{0}$ and covariance $\mathbf{V}$. We then put $\mathbf{t} = \boldsymbol{\mu} + \mathbf{LN}$ resulting in a vector $\mathbf{t}$ with the desired distribution. The calculations for a single simulated Treasury yield curve would look like this:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ T_7 \\ T_8 \end{bmatrix} = \begin{bmatrix} 1.22 \\ 1.43 \\ 1.66 \\ 2.15 \\ 2.56 \\ 2.93 \\ 3.49 \\ 3.67 \end{bmatrix} + \begin{bmatrix} 1.63 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 1.48 & 0.15 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 1.36 & 0.26 & 0.07 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 1.14 & 0.36 & 0.20 & 0.12 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.97 & 0.39 & 0.27 & 0.20 & 0.06 & 0.00 & 0.00 & 0.00 \\ 0.82 & 0.35 & 0.32 & 0.30 & 0.13 & 0.08 & 0.00 & 0.00 \\ 0.66 & 0.34 & 0.38 & 0.41 & 0.22 & 0.16 & 0.07 & 0.00 \\ 0.53 & 0.27 & 0.33 & 0.39 & 0.22 & 0.11 & 0.02 & 0.09 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \\ N_5 \\ N_6 \\ N_7 \\ N_8 \end{bmatrix}$$

Using PCA. First we observe that, in Figure 1, we may delete columns $k+1$ through p of $\widetilde{\mathbf{Y}}$ and rows $k+1$ through p of $\mathbf{Q}^T$ without changing the result:

Figure 3.



With $k = 3$, we may simulate Treasury yield curves with approximately the desired multivariate normal distribution as follows. First, we generate multivariate mean-zero normals $\mathbf{y} = (Y_1, Y_2, Y_3)^T$ with covariance matrix

$$\begin{pmatrix} 11.04 & 0 & 0 \\ 0 & 0.97 & 0 \\ 0 & 0 & 0.08 \end{pmatrix},$$

the covariance matrix for the first three columns of the $\mathbf{Y}$ matrix. That is, take

$$\mathbf{y} = (Y_1, Y_2, Y_3)^T = (\sqrt{11.04}\, N_1, \sqrt{0.97}\, N_2, \sqrt{0.08}\, N_3)^T,$$

where N_1, N_2, and N_3 are independent standard normals. Then put $\mathbf{t}^T = \boldsymbol{\mu}^T + \mathbf{y}^T\widehat{\mathbf{Q}}^T$, where

$$\widehat{\mathbf{Q}}^T = \boxed{\begin{matrix} \text{——} \ \mathbf{q}_1^T \ \text{——} \\ \text{——} \ \mathbf{q}_2^T \ \text{——} \\ \text{——} \ \mathbf{q}_3^T \ \text{——} \end{matrix}}_{\,(3\times 8)}.$$

Here, of course, $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ are the eigenvectors corresponding to the three eigenvalues 11.04, 0.97, and 0.08. Stated in terms of column vectors, we have $\mathbf{t} = \boldsymbol{\mu} + \widehat{\mathbf{Q}}\mathbf{y}$, where $\widehat{\mathbf{Q}}$ is a $p \times k = 8 \times 3$ matrix whose columns are the first three eigenvectors.

A slightly more efficient way to accomplish this is to throw the scaling factors $\sqrt{11.04}$, $\sqrt{0.97}$, and $\sqrt{0.08}$ present in the $\mathbf{y}$ vector on to the eigenvectors in $\widehat{\mathbf{Q}}$. For each simulation we then generate standard independent normals N_1, N_2, N_3 and put:

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ T_7 \\ T_8 \end{bmatrix} = \begin{bmatrix} 1.22 \\ 1.43 \\ 1.66 \\ 2.15 \\ 2.56 \\ 2.93 \\ 3.49 \\ 3.67 \end{bmatrix} + \begin{bmatrix} 1.56 & -0.41 & 0.15 \\ 1.46 & -0.31 & 0.02 \\ 1.36 & -0.19 & -0.07 \\ 1.20 & 0.05 & -0.13 \\ 1.06 & 0.21 & -0.12 \\ 0.93 & 0.34 & -0.02 \\ 0.83 & 0.51 & 0.08 \\ 0.66 & 0.48 & 0.10 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix}.$$

Column 1 above is $\sqrt{11.04}\,\mathbf{q}_1$, column 2 is $\sqrt{0.97}\,\mathbf{q}_2$, and column 3 is $\sqrt{0.08}\,\mathbf{q}_3$. While this approach yields a vector $\mathbf{t}$ whose distribution is only approximately what is desired, the calculations involved are obviously more efficient than the full-blown Cholesky approach.

Accompanying Code

`Fact1.cpp` illustrates that $\mathbf{VQ} = \mathbf{Q\Lambda}$ (Fact 1, above) holds for a specific choice of $\mathbf{V}$. `TreasuryPCAnalysis.cpp` performs the Treasury PCA as described above. For $k = 0, 1, \ldots, 8$, quarterly yield curve fits from February 2006 to February 2018 may be viewed using the TeX file `TreasuryPCA.tex`. The TeX file `SpecificYieldCurve.tex` displays the $k = 0, 1, 2$ and 3 fits for a specific date specified by the user.