Chapter 1

Vectors and matrices.

Column vectors and row vectors. Column form and row form of a matrix.

Column-based and row-based matrix–vector and matrix–matrix multiplication.

Covariance matrix computation from time series data.

Matrix rank. Nullspace and range of a matrix. Linear independence.

A one period market model example.

Nonsingular matrices and the inverse of a matrix.

Diagonal matrices. Matrix multiplication by diagonal matrices.

Converting between covariance and correlation matrices.

Correlation matrix computation from time series data.

Lower and upper triangular matrices. Tridiagonal matrices.

1.1 Column and row vectors. Column form and row form of a matrix.

An n–dimensional vector $v \in \mathbb{R}^n$ is denoted by $v = (v_i)_{i=1:n}$ and has n components $v_i \in \mathbb{R}$, for $i = 1 : n$.[1]

The vector $v = (v_i)_{i=1:n}$ is a **column vector** of size n if

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{1.1}$$

A column vector is also called an $n \times 1$ vector.

[1] The vectors and matrices considered here have entries which are real numbers. While complex numbers will occur naturally (for example, eigenvalues of a matrices with real entries may be complex numbers), the presentation and the notations in this book will be specific to vectors and matrices with real entries.

The vector v^t is a **row vector**[2] of size n if

$$v^t \;=\; (v_1\ v_2\ \ldots\ v_n). \tag{1.2}$$

A row vector is also called an $1 \times n$ vector.

Unless otherwise specified, a vector v denoted by $v = (v_i)_{i=1:n}$ is a column vector.

An $m \times n$ matrix $A = (A(j,k))_{j=1:m,k=1:n}$ is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, i.e., $A : \mathbb{R}^n \to \mathbb{R}^m$. The matrix A has m rows and n columns. Rather than using the entry by entry notation above for the matrix A, we will use either a column–based notation (more often), or a row–based notation, both being better suited for numerical computations.

The **column form** of the matrix A is

$$A \;=\; (a_1 \mid a_2 \mid \ldots \mid a_n) \;=\; \mathrm{col}\,(a_k)_{k=1:n}\,, \tag{1.3}$$

where a_k is the k-th column[3] of A, $k = 1 : n$.

The **row form** of the matrix A is

$$A \;=\; \begin{pmatrix} r_1 \\ -- \\ r_2 \\ -- \\ \vdots \\ -- \\ r_m \end{pmatrix} \;=\; \mathrm{row}\,(r_j)_{j=1:m}\,, \tag{1.4}$$

where r_j is the j-th row[4] of A, $j = 1 : m$.

Row Vector – Column Vector multiplication:[5]
Let $v = (v_i)_{i=1:n}$ be a column vector of size n, and let $w^t = (w_1\ w_2\ \ldots\ w_n)$ be a row vector of size n. Then,

$$w^t v \;=\; \sum_{i=1}^{n} w_i v_i. \tag{1.5}$$

Column Vector – Row Vector multiplication:
Let $v = (v_j)_{j=1:m}$ be a column vector of size m, and let $w^t = (w_1\ w_2\ \ldots\ w_n)$ be a row vector of size n. Then, vw^t is an $m \times n$ matrix with the following entries:

$$(vw^t)(j,k) \;=\; v_j w_k, \;\; \forall\ j = 1 : m,\ k = 1 : n. \tag{1.6}$$

Matrix – Column Vector multiplication:
Let $A = \mathrm{col}\,(a_k)_{k=1:n}$ be an $m \times n$ matrix given by the column form (1.3), and let

[2]The notation v^t emphasizes the fact that a row vector is the transpose of a column vector; see also Definition 1.1.

[3]For every $k = 1 : n$, the column vector a_k is given by $a_k = (A(j,k))_{j=1:m}$.

[4]For every $j = 1 : m$, the row vector r_j is given by $r_j = (A(j,k))_{k=1:n}$.

[5]Formula (1.5) is the same as formula (5.2) for the Euclidean inner product of two vectors.

$v = (v_k)_{k=1:n}$ be a column vector of size n given by (1.1). Then,

$$Av \ = \ \sum_{k=1}^{n} v_k a_k. \tag{1.7}$$

In other words, the result of the multiplication of the column vector v by the matrix A is a column vector Av which is the linear combination[6] of the columns of A with coefficients equal to the corresponding entries of v.

If $A = \mathrm{row}\,(r_j)_{j=1:m}$ is the row form of A, then the j–th entry of the $m \times 1$ column vector Av is

$$(Av)(j) \ = \ r_j v, \quad \forall\ 1 \le j \le m. \tag{1.8}$$

Row Vector – Matrix multiplication:
Let $A = \mathrm{row}\,(r_j)_{j=1:m}$ be an $m \times n$ matrix given by the row form (1.4), and let $w^t = (w_1\ w_2\ \ldots\ w_m)$ be a row vector of size m. Then,

$$w^t A \ = \ \sum_{j=1}^{m} w_j r_j. \tag{1.9}$$

In other words, the result of the multiplication of the row vector w^t by the matrix A (from the right) is a row vector $w^t A$ which is the linear combination of the rows of A with coefficients equal to the corresponding entries of w^t.

If $A = \mathrm{col}\,(a_k)_{k=1:n}$ is the column form of A, then the k–th entry of the $1 \times n$ row vector $w^t A$ is

$$(w^t A)(k) \ = \ w^t a_k, \quad \forall\ 1 \le k \le n. \tag{1.10}$$

Matrix – Matrix multiplication:
(i) Let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix given by $B = \mathrm{col}\,(b_k)_{k=1:p}$. Then, AB is the $m \times p$ matrix given by

$$AB \ = \ \mathrm{col}\,(Ab_k)_{k=1:p} \ = \ (Ab_1 \mid Ab_2 \mid \ \ldots \ \mid Ab_p)\,. \tag{1.11}$$

The result of multiplying the matrices A and B is a matrix whose columns are the columns of B multiplied by the matrix A.

(ii) Let A be an $m \times n$ matrix given by $A = \mathrm{row}\,(r_j)_{j=1:m}$, and let B be an $n \times p$ matrix. Then, AB is the $m \times p$ matrix given by

$$AB \ = \ \mathrm{row}\,(r_j B)_{j=1:m} \ = \ \begin{pmatrix} r_1 B \\ -- \\ r_2 B \\ -- \\ \vdots \\ -- \\ r_m B \end{pmatrix}. \tag{1.12}$$

[6] A linear combination of the vectors $w_1, w_2, \ldots, w_n$ is any sum of these vectors multiplied by real coefficients, i.e., $c_1 w_1 + c_2 w_2 + \ldots + c_n w_n$, where $c_i \in \mathbb{R}$, $i = 1 : n$; see also Definition 1.5.

The result of multiplying the matrices A and B is a matrix whose rows are the rows of A multiplied by the matrix B.

(iii) Let A be an $m \times n$ matrix given by $A = \mathrm{row}\,(r_j)_{j=1:m}$, and let B be an $n \times p$ matrix given by $B = \mathrm{col}\,(b_k)_{k=1:p}$. Then, AB is the $m \times p$ matrix whose entries are given by

$$(AB)(j,k) \;=\; r_j b_k, \quad \forall\, j = 1:m,\; k = 1:p. \tag{1.13}$$

Note that, since r_j is a $1 \times n$ row vector and b_k is a $n \times 1$ column vector, it follows from (1.5) that the multiplication from (1.13) can be performed.

Matrix – Matrix – Matrix multiplication:
Let A be an $m \times n$ matrix given by $A = \mathrm{row}\,(r_j)_{j=1:m}$, let B be an $n \times p$ matrix, and let C be a $p \times l$ matrix given by $C = \mathrm{col}\,(c_k)_{k=1:l}$. Then, ABC is the $m \times l$ matrix whose entries are given by

$$(ABC)(j,k) \;=\; r_j B c_k, \quad \forall\, j = 1:m,\; k = 1:l. \tag{1.14}$$

Note that (1.14) follows from (1.13), since $BC = \mathrm{col}\,(Bc_k)_{k=1:l}$; cf. (1.11).

Note that matrix multiplication is associative, i.e., $ABC = (AB)C = A(BC)$.

We emphasize again that we almost exclusively think of a matrix as either a collection of column vectors, or as a collection of row vectors, rather than as a collection of individual entries. For numerical purposes, this is an efficient way to implement matrices. Also, linear algebra proofs using the column form or the row form of a matrix are more insightful and more compact than proofs using individual entries of a matrix. Most of the proofs from this book use a vector–based approach.

Definition 1.1. *The transpose of an $n \times 1$ column vector $v = (v_i)_{i=1:n}$ is the $1 \times n$ row vector $v^t = (v_1\ v_2\ \ldots\ v_n)$. The transpose of an $1 \times n$ row vector $r = (r_1\ r_2\ \ldots\ r_n)$ is the $n \times 1$ column vector $r^t = (r_i)_{i=1:n}$.*

Note that

$$(cv)^t \;=\; cv^t, \quad \forall\, v \in \mathbb{R}^n,\; c \in \mathbb{R}. \tag{1.15}$$

Definition 1.2. *The transpose matrix A^t of an $m \times n$ matrix A is an $n \times m$ matrix given by*

$$A^t(k,j) \;=\; A(j,k), \quad \forall\, k = 1:n,\; j = 1:m. \tag{1.16}$$

Transposing a matrix switches the column form of the matrix to a row form, and the row form of the matrix to a column form as follows:

$$A = \mathrm{col}\,(a_k)_{k=1:n} \quad\Longleftrightarrow\quad A^t = \mathrm{row}\left(a_k^t\right)_{k=1:n}; \tag{1.17}$$

$$A = \mathrm{row}\,(r_j)_{j=1:m} \quad\Longleftrightarrow\quad A^t = \mathrm{col}\left(r_j^t\right)_{j=1:m}. \tag{1.18}$$

From (1.16), we find that, for any matrix A,

$$(A^t)^t \;=\; A, \tag{1.19}$$

and, for any matrices A and B of the same size,

$$(A+B)^t \;=\; A^t + B^t. \tag{1.20}$$

Lemma 1.1. *Let A be an $m \times n$ matrix and let v be a column vector of size n. Then,*

$$(Av)^t \;=\; v^t A^t. \tag{1.21}$$

Proof. Let $A = \mathrm{col}\,(a_k)_{k=1:n}$ and $v = (v_i)_{i=1:n}$. Then, $Av = \sum_{i=1}^n v_i a_i$, and

$$(Av)^t \;=\; \left(\sum_{i=1}^n v_i a_i\right)^t \;=\; \sum_{i=1}^n (v_i a_i)^t \;=\; \sum_{i=1}^n v_i a_i^t, \tag{1.22}$$

since $v_i \in \mathbb{R}$; see (1.15).

Note that $A^t = \mathrm{row}\,\left(a_k^t\right)_{k=1:n}$; cf. (1.17). Then, from (1.9), it follows that

$$v^t A^t \;=\; \sum_{i=1}^n v_i a_i^t. \tag{1.23}$$

From (1.22) and (1.23), we conclude that $(Av)^t = v^t A^t$. □

It is very important to note that the transpose of the product of two matrices is *not* the product of the transposes of the two matrices,[7] i.e., $(AB)^t \neq A^t B^t$. Instead, the following result holds:[8]

Lemma 1.2. *Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Then,*

$$(AB)^t \;=\; B^t A^t. \tag{1.24}$$

Proof. Recall from (1.11) that, if $B = \mathrm{col}\,(b_k)_{k=1:p}$, then $AB = \mathrm{col}\,(Ab_k)_{k=1:p}$. Thus, from (1.17), we obtain that

$$(AB)^t \;=\; \left(\mathrm{col}\,(Ab_k)_{k=1:p}\right)^t \;=\; \mathrm{row}\,\left((Ab_k)^t\right)_{k=1:p}.$$

Using (1.21), (1.12), and the fact that $B^t = \mathrm{row}\,\left(b_k^t\right)_{k=1:p}$, see (1.17) we conclude that

$$(AB)^t \;=\; \mathrm{row}\,\left(b_k^t A^t\right)_{k=1:p} \;=\; \left(\mathrm{row}\,\left(b_k^t\right)_{k=1:p}\right) A^t \;=\; B^t A^t.$$

□

Definition 1.3. *A matrix with the same number of rows and columns is called a square matrix.*

Note that an $n \times n$ square matrix is also called a square matrix of size n.

Definition 1.4. *A square matrix is symmetric if and only if the matrix and its transpose are the same. In other words, a square matrix A of size n is symmetric if and only if $A = A^t$, i.e.,*

$$A(j,k) = A(k,j), \quad \forall\ 1 \le j < k \le n;$$

[7] A similar property holds for inverses of matrices, i.e., $(AB)^{-1} \neq A^{-1}B^{-1}$. Moreover, $(AB)^{-1} = B^{-1}A^{-1}$; see Lemma 1.7 for details.

[8] The result of Lemma 1.2 extends as follows: $\left(\prod_{i=1}^p A_i\right)^t = \prod_{i=1}^p A_{p+1-i}^t$. A proof can be given by induction; see an exercise at the end of this chapter.

The product of two symmetric matrices is not necessarily a symmetric matrix, as seen in the example below.

Example: Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$ be two symmetric matrices. Then,

$$AB = \begin{pmatrix} 4 & 1 \\ 5 & 2 \end{pmatrix} \neq (AB)^t = \begin{pmatrix} 4 & 5 \\ 1 & 2 \end{pmatrix} \quad \square$$

The identity matrix,[9] denoted by I, is a square matrix with entries equal to 1 on the main diagonal and equal to 0 everywhere else, i.e.,

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

The k-th column of the identity matrix is denoted by e_k. Thus,

$$e_k(i) = 0, \text{ for } 1 \leq i \neq k \leq n \quad \text{and} \quad e_k(k) = 1. \tag{1.25}$$

The column form and row form of the identity matrix I are, respectively,

$$I = \text{col}\,(e_k)_{k=1:n}\,; \quad I = \text{row}\,\left(e_k^t\right)_{k=1:n}\,;$$

cf. (1.17), since $I = I^t$.

Lemma 1.3. *(i) Let $A = col\,(a_k)_{k=1:n}$ be an $m \times n$ matrix. If e_k is the k-th column of the $n \times n$ identity matrix, then*

$$Ae_k = a_k, \quad \forall\ k = 1:n, \tag{1.26}$$

and therefore $AI = A$.

(ii) Let $A = row\,(r_j)_{j=1:m}$ be an $m \times n$ matrix. If e_j is the j-th column of the $m \times m$ identity matrix, then

$$e_j^t A = r_j, \quad \forall\ j = 1:m, \tag{1.27}$$

and therefore $IA = A$.

Proof. (i) Let $A = \text{col}\,(a_k)_{k=1:n}$. Recall from (1.25) that $e_k(k) = 1$ and $e_k(i) = 0$, for $i \neq k$. From (1.7), we obtain that

$$Ae_k = \sum_{i=1}^{n} e_k(i) a_i = a_k. \tag{1.28}$$

If $I = \text{col}\,(e_k)_{k=1:n}$, it follows from (1.11) and (1.28) that

$$AI = \text{col}\,(Ae_k)_{k=1:n} = \text{col}\,(a_k)_{k=1:n} = A.$$

[9]The $n \times n$ identity matrix is sometimes denoted by I_n. We do not use this notation, but denote by I identity matrices of any size.

(ii) Let $A = \text{row}\,(r_j)_{j=1:m}$. Recall from (1.25) that $e_j(j) = 1$ and $e_j(i) = 0$, for $i \neq j$. From (1.9), we find that

$$e_j^t A \;=\; \sum_{i=1}^{m} e_j(i) r_i \;=\; r_j. \tag{1.29}$$

If $I = \text{row}\left(e_j^t\right)_{j=1:m}$, it follows from (1.12) and (1.29) that

$$IA \;=\; \text{row}\left(e_j^t A\right)_{j=1:m} \;=\; \text{row}\,(r_j)_{j=1:m} \;=\; A.$$

□

1.1.1 Covariance matrix computation from time series data

Let $X_1, X_2, \ldots, X_n$ be random variables given by time series data at N data points t_i, $i = 1 : N$. In other words, the values of $X_k(t_i)$ are given for all $k = 1 : n$ and $i = 1 : N$.

Denote by $\widehat{\mu}_{X_k}$ the sample mean of the random variable X_k, for $k = 1 : n$, i.e.,

$$\widehat{\mu}_{X_k} \;=\; \frac{1}{N} \sum_{i=1}^{N} X_k(t_i).$$

The sample covariance matrix $\widehat{\Sigma}_{\mathbf{X}}$ of the n random variables $X_1, X_2, \ldots, X_n$ is the $n \times n$ square matrix with entries

$$\widehat{\Sigma}_{\mathbf{X}}(j,k) \;=\; \widehat{\text{cov}}(X_j, X_k), \quad \forall\; 1 \leq j,k \leq n, \tag{1.30}$$

where $\widehat{\text{cov}}(X_j, X_k)$ is the unbiased sample covariance of the random variables X_j and X_k given by

$$\widehat{\text{cov}}(X_j, X_k) \;=\; \frac{1}{N-1} \sum_{i=1}^{N} (X_j(t_i) - \widehat{\mu}_{X_j})(X_k(t_i) - \widehat{\mu}_{X_k}). \tag{1.31}$$

From (1.30) and (1.31), we find that

$$\widehat{\Sigma}_{\mathbf{X}}(j,k) \;=\; \frac{1}{N-1} \sum_{i=1}^{N} (X_j(t_i) - \widehat{\mu}_{X_j})(X_k(t_i) - \widehat{\mu}_{X_k}). \tag{1.32}$$

The sample covariance matrix $\widehat{\Sigma}_{\mathbf{X}}$ is symmetric since, from (1.32), it follows that

$$\begin{aligned}
\widehat{\Sigma}_{\mathbf{X}}(j,k) \;&=\; \frac{1}{N-1} \sum_{i=1}^{N} (X_j(t_i) - \widehat{\mu}_{X_j})(X_k(t_i) - \widehat{\mu}_{X_k}) \\
&=\; \frac{1}{N-1} \sum_{i=1}^{N} (X_k(t_i) - \widehat{\mu}_{X_k})(X_j(t_i) - \widehat{\mu}_{X_j}) \\
&=\; \widehat{\Sigma}_{\mathbf{X}}(k,j), \quad \forall\; 1 \leq j,k \leq n.
\end{aligned}$$

The sample covariance matrix can be computed efficiently by using matrix formulation for the time series data $X_k(t_i)$, $i = 1 : N$, $k = 1 : n$, as shown below.

Let $T_{\mathbf{X}}$ be the corresponding $N \times n$ matrix of time series data, i.e., let $T_{\mathbf{X}} = (T_{\mathbf{X}}(i,k))_{i=1:N,k=1:n}$ with

$$T_{\mathbf{X}}(i,k) = X_k(t_i), \quad \forall\ 1 \le k \le n,\ 1 \le i \le N. \tag{1.33}$$

Let $\overline{T}_{\mathbf{X}}$ be the $N \times n$ matrix of time series data where the sample mean of each random variable is subtracted from the corresponding time series data, i.e., let $\overline{T}_{\mathbf{X}} = (\overline{T}_{\mathbf{X}}(i,k))_{i=1:N,k=1:n}$ with

$$\overline{T}_{\mathbf{X}}(i,k) \ = \ X_k(t_i) - \widehat{\mu}_{X_k}, \quad \forall\ 1 \le k \le n,\ 1 \le i \le N. \tag{1.34}$$

Then, the sample covariance matrix $\widehat{\Sigma}_{\mathbf{X}}$ can be computed from $\overline{T}_{\mathbf{X}}$ as follows:

$$\widehat{\Sigma}_{\mathbf{X}} \ = \ \frac{1}{N-1}\, \overline{T}_{\mathbf{X}}^t \overline{T}_{\mathbf{X}}. \tag{1.35}$$

For clarity, we include an example below and the proof of (1.35).

Example: The end of day adjusted close prices for Apple, Facebook, Google, Microsoft, and Yahoo between 1/10/2013 and 1/29/2013 were:

Date	AAPL	FB	GOOG	MSFT	YHOO
1/10/2013	523.51	31.30	741.48	26.46	18.99
1/11/2013	520.30	31.72	739.99	26.83	19.29
1/14/2013	501.75	30.95	723.25	26.89	19.43
1/15/2013	485.92	30.10	724.93	27.21	19.52
1/16/2013	506.09	29.85	715.19	27.04	20.07
1/17/2013	502.68	30.14	711.32	27.25	20.13
1/18/2013	500.00	29.66	704.51	27.25	20.02
1/22/2013	504.77	30.73	702.87	27.15	19.90
1/23/2013	514.01	30.82	741.50	27.61	20.11
1/24/2013	450.50	31.08	754.21	27.63	20.44
1/25/2013	439.88	31.54	753.67	27.88	20.37
1/28/2013	449.83	32.47	750.73	27.91	20.31
1/29/2013	458.27	30.79	753.68	28.01	19.70

The time series matrix of the daily returns[10] of the five stocks above between 1/11/2013 and 1/29/2013 is

$$T_{\mathbf{X}} \ = \ \begin{pmatrix} -0.0061 & 0.0134 & -0.0020 & 0.0140 & 0.0158 \\ -0.0357 & -0.0243 & -0.0226 & 0.0022 & 0.0073 \\ -0.0315 & -0.0275 & 0.0023 & 0.0119 & 0.0046 \\ 0.0415 & -0.0083 & -0.0134 & -0.0062 & 0.0282 \\ -0.0067 & 0.0097 & -0.0054 & 0.0078 & 0.0030 \\ -0.0053 & -0.0159 & -0.0096 & 0.0000 & -0.0055 \\ 0.0095 & 0.0361 & -0.0023 & -0.0037 & -0.0060 \\ 0.0183 & 0.0029 & 0.0550 & 0.0169 & 0.0106 \\ -0.1236 & 0.0084 & 0.0171 & 0.0007 & 0.0164 \\ -0.0236 & 0.0148 & -0.0007 & 0.0090 & -0.0034 \\ 0.0226 & 0.0295 & -0.0039 & 0.0011 & -0.0029 \\ 0.0188 & -0.0517 & 0.0039 & 0.0036 & -0.0300 \end{pmatrix},$$

[10] Unless specified otherwise, the return between times τ_1 and τ_2 of an asset with spot prices $S(\tau_1)$ and $S(\tau_2)$ will mean the percentage return, which is $\frac{S(\tau_2)-S(\tau_1)}{S(\tau_1)}$.

where, e.g., the daily return of GOOG on 1/24/2013 is

$$\frac{754.21 - 741.50}{741.50} \;=\; 0.0171 \;=\; T_{\mathbf{X}}(9,3),$$

and the daily return of GOOG on 1/28/2013 is

$$\frac{750.73 - 753.67}{753.67} \;=\; -\,0.0039 \;=\; T_{\mathbf{X}}(11,3).$$

The sample means of the returns of the five stocks are -0.0101 (AAPL), -0.0011 (FB), 0.0015 (GOOG), 0.0048 (MSFT), and 0.0032 (YHOO). By subtracting the sample mean of each column of $T_{\mathbf{X}}$ we obtain from (1.34) that

$$\overline{T}_{\mathbf{X}} \;=\; \begin{pmatrix} 0.0040 & 0.0145 & -0.0035 & 0.0092 & 0.0126 \\ -0.0255 & -0.0232 & -0.0242 & -0.0025 & 0.0041 \\ -0.0214 & -0.0264 & 0.0008 & 0.0071 & 0.0015 \\ 0.0517 & -0.0072 & -0.0150 & -0.0110 & 0.0250 \\ 0.0034 & 0.0108 & -0.0069 & 0.0030 & -0.0002 \\ 0.0048 & -0.0149 & -0.0111 & -0.0048 & -0.0086 \\ 0.0197 & 0.0371 & -0.0039 & -0.0084 & -0.0092 \\ 0.0285 & 0.0040 & 0.0534 & 0.0122 & 0.0074 \\ -0.1134 & 0.0095 & 0.0156 & -0.0041 & 0.0132 \\ -0.0134 & 0.0159 & -0.0022 & 0.0043 & -0.0066 \\ 0.0328 & 0.0306 & -0.0054 & -0.0037 & -0.0061 \\ 0.0289 & -0.0507 & 0.0024 & -0.0012 & -0.0332 \end{pmatrix}. \qquad (1.36)$$

We now show that the formula (1.35) holds, i.e.,

$$\widehat{\Sigma}_{\mathbf{X}} \;=\; \frac{1}{N-1}\,\overline{T}_{\mathbf{X}}^t\,\overline{T}_{\mathbf{X}}; \qquad (1.37)$$

see also Theorem 7.1 and the proof therein.

From (1.34), we find that, for any $1 \le j,k \le n$,

$$\overline{T}_{\mathbf{X}}(i,k) \;=\; X_k(t_i) - \widehat{\mu}_{X_k} \quad \text{and} \quad \overline{T}_{\mathbf{X}}(i,j) \;=\; X_j(t_i) - \widehat{\mu}_{X_j}, \quad \forall\; i = 1:N. \qquad (1.38)$$

Then, from (1.32) and (1.38), it follows that

$$\begin{aligned} \widehat{\Sigma}_{\mathbf{X}}(j,k) \;&=\; \frac{1}{N-1}\sum_{i=1}^{N}(X_j(t_i) - \widehat{\mu}_{X_j})(X_k(t_i) - \widehat{\mu}_{X_k}) && (1.39)\\ &=\; \frac{1}{N-1}\sum_{i=1}^{N}\overline{T}_{\mathbf{X}}(i,j)\overline{T}_{\mathbf{X}}(i,k), \quad \forall\; 1 \le j,k \le n. && (1.40) \end{aligned}$$

Let $\overline{T}_{X_k}$ be the $N \times 1$ column vector of the time series data for the random variable X_k with $\widehat{\mu}_{X_k}$ subtracted from each data value, i.e.,

$$\overline{T}_{X_k} \;=\; \left(X_k(t_i) - \widehat{\mu}_{X_k}\right)_{i=1:N}.$$

The time series matrix $\overline{T}_{\mathbf{X}} = (\overline{T}_{\mathbf{X}}(i,k))_{i=1:N,k=1:n}$ has the following column form:

$$\overline{T}_{\mathbf{X}} \;=\; \mathrm{col}\left(\overline{T}_{X_k}\right)_{k=1:n}.$$

Moreover,

$$\overline{T}_{\mathbf{X}}(i,j) = \overline{T}_{X_j}(i) \quad \text{and} \quad \overline{T}_{\mathbf{X}}(i,k) = \overline{T}_{X_k}(i), \quad \forall\ 1 \le j,k \le n,\ 1 \le i \le N,$$

and, from (1.40), we obtain that

$$\widehat{\Sigma}_{\mathbf{X}}(j,k) = \frac{1}{N-1} \sum_{i=1}^{N} \overline{T}_{X_j}(i)\overline{T}_{X_k}(i) \tag{1.41}$$

$$= \frac{1}{N-1}\ \overline{T}^t_{X_j}\overline{T}_{X_k}, \quad \forall\ 1 \le j,k \le n, \tag{1.42}$$

where the last equality follows from the row vector–column vector multiplication formula (1.5).

Since $\overline{T}_{\mathbf{X}} = \operatorname{col}\left(\overline{T}_{X_k}\right)_{k=1:n}$, it follows that $\overline{T}^t_{\mathbf{X}} = \operatorname{row}\left(\overline{T}^t_{X_j}\right)_{j=1:n}$; see (1.17). From (1.13), we obtain that the (j,k) entry of the matrix $\overline{T}^t_{\mathbf{X}}\overline{T}_{\mathbf{X}}$ is

$$(\overline{T}^t_{\mathbf{X}}\overline{T}_{\mathbf{X}})(j,k) = \overline{T}^t_{X_j}\overline{T}_{X_k}, \quad \forall\ 1 \le j,k \le n. \tag{1.43}$$

Then, from (1.42) and (1.43), we conclude that

$$\widehat{\Sigma}_{\mathbf{X}}(j,k) = \frac{1}{N-1}\ (\overline{T}^t_{\mathbf{X}}\overline{T}_{\mathbf{X}})(j,k), \quad \forall\ 1 \le j,k \le n,$$

and therefore

$$\widehat{\Sigma}_{\mathbf{X}} = \frac{1}{N-1}\ \overline{T}^t_{\mathbf{X}}\overline{T}_{\mathbf{X}},$$

which is what we wanted to prove; see (1.37).

Example (continued):

The sample covariance matrix $\widehat{\Sigma}_{\mathbf{X}}$ of the daily returns of AAPL, FB, GOOG, MSFT, YHOO between 1/11/2013 and 1/29/2013 can be computed using formula (1.35) with $N = 12$ and $\overline{T}_{\mathbf{X}}$ given by (1.36). We find that

$$\widehat{\Sigma}_{\mathbf{X}} = \begin{pmatrix} 0.0018 & 0.0000 & -0.0001 & 0.0000 & -0.0001 \\ 0.0000 & 0.0006 & 0.0001 & 0.0000 & 0.0001 \\ -0.0001 & 0.0001 & 0.0004 & 0.0001 & 0.0000 \\ 0.0000 & 0.0000 & 0.0001 & 0.0001 & 0.0000 \\ -0.0001 & 0.0001 & 0.0000 & 0.0000 & 0.0002 \end{pmatrix}. \quad \square \tag{1.44}$$

More properties of covariance matrices obtained from time series data can be found in section 7.2.

1.2 Matrix rank, nullspace, and range of a matrix

Definition 1.5. *Let w_1, w_2, ..., w_p be vectors of the same size. The vectors w_1, w_2, ..., w_p are linearly independent if and only if the only linear combination of these vectors that is equal to 0 has all coefficients equal to 0, i.e.,*

$$\text{if} \quad \sum_{i=1}^{p} c_i w_i = 0, \quad \text{with } c_i \in \mathbb{R}, i = 1:p, \quad \text{then} \quad c_i = 0,\ \forall\ i = 1:p.$$