APPENDIX

Linear Algebra Basics

B

B.1 POSITIVE DEFINITE AND SYMMETRIC MATRICES

- An $l \times l$ real matrix A is called *positive definite* if for *every* nonzero vector $\boldsymbol{x}$ the following is true:

$$\boldsymbol{x}^T A \boldsymbol{x} > 0 \tag{B.1}$$

If equality with zero is allowed, A is called *nonnegative or positive semidefinite*.

- It is easy to show that all eigenvalues of such a matrix are positive. Indeed, let λ_i be one eigenvalue and $\boldsymbol{v}_i$ the corresponding unit norm eigenvector ($\boldsymbol{v}_i^T \boldsymbol{v}_i = 1$). Then by the respective definitions

$$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \quad \text{or} \tag{B.2}$$

$$0 < \boldsymbol{v}_i^T A \boldsymbol{v}_i = \lambda_i \tag{B.3}$$

Since the determinant of a matrix is equal to the product of its eigenvalues, we conclude that the determinant of a positive definite matrix is also positive.

- Let A be an $l \times l$ symmetric matrix, $A^T = A$. Then the eigenvectors corresponding to distinct eigenvalues are orthogonal. Indeed, let $\lambda_i \neq \lambda_j$ be two such eigenvalues. From the definitions we have

$$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \tag{B.4}$$

$$A\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j \tag{B.5}$$

Multiplying (B.4) on the left by $\boldsymbol{v}_j^T$ and the transpose of (B.5) on the right by $\boldsymbol{v}_i$, we obtain

$$\boldsymbol{v}_j^T A \boldsymbol{v}_i - \boldsymbol{v}_j^T A \boldsymbol{v}_i = 0 = (\lambda_i - \lambda_j)\boldsymbol{v}_j^T \boldsymbol{v}_i \tag{B.6}$$

Thus, $\boldsymbol{v}_j^T \boldsymbol{v}_i = 0$. Furthermore, it can be shown that even if the eigenvalues are not distinct, we can still find a set of orthogonal eigenvectors. The same is true for Hermitian matrices, in case we deal with more general complex-valued matrices.

- Based on this, it is now straightforward to show that a symmetric matrix A can be diagonalized by the similarity transformation

$$\Phi^T A \Phi = \Lambda \tag{B.7}$$

where matrix Φ has as its columns the unit eigenvectors ($\boldsymbol{v}_i^T \boldsymbol{v}_i = 1$) of A, that is,

$$\Phi = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l] \tag{B.8}$$

and Λ is the diagonal matrix with elements the corresponding eigenvalues of A. From the orthonormality of the eigenvectors it is obvious that $\Phi^T \Phi = I$, that is, Φ is a unitary matrix, $\Phi^T = \Phi^{-1}$. The proof is similar for Hermitian complex matrices as well.

B.2 CORRELATION MATRIX DIAGONALIZATION

Let $\boldsymbol{x}$ be a random vector in the l-dimensional space. Its correlation matrix is defined as $R = E[\boldsymbol{x}\boldsymbol{x}^T]$. Matrix R is readily seen to be positive semidefinite. For our purposes we will assume that it is positive definite, thus invertible. Moreover, it is symmetric, and hence it can always be diagonalized

$$\Phi^T R \Phi = \Lambda \tag{B.9}$$

where Φ is the matrix consisting of the (orthogonal) eigenvectors and Λ the diagonal matrix with the corresponding eigenvalues on its diagonal. Thus, we can always transform $\boldsymbol{x}$ into another random vector whose elements are uncorrelated. Indeed

$$\boldsymbol{x}_1 \equiv \Phi^T \boldsymbol{x} \tag{B.10}$$

Then the new correlation matrix is $R_1 = \Phi^T R \Phi = \Lambda$. Furthermore, if $\Lambda^{1/2}$ is the diagonal matrix whose elements are the square roots of the eigenvalues of R ($\Lambda^{1/2}\Lambda^{1/2} = \Lambda$), then it is readily shown that the transformed random vector

$$\boldsymbol{x}_1 \equiv \Lambda^{-1/2} \Phi^T \boldsymbol{x} \tag{B.11}$$

has uncorrelated elements with unit variance. $\Lambda^{-1/2}$ denotes the inverse of $\Lambda^{1/2}$. It is now easy to see that if the correlation matrix of a random vector is the identity matrix I, then this is invariant under any unitary transformation $A^T \boldsymbol{x}, A^T A = I$. That is, the transformed variables are also uncorrelated with unit variance. A useful by-product of this is the following lemma.

Lemma *Let $\boldsymbol{x}, \boldsymbol{y}$ be two random vectors with correlation matrices R_x, R_y, respectively. Then there is a linear transformation that diagonalizes both matrices simultaneously.*

Proof. Let Φ be the eigenvector matrix diagonalizing R_x. Then the transformation

$$\boldsymbol{x}_1 \equiv \Lambda^{-1/2}\Phi^T \boldsymbol{x} \tag{B.12}$$

$$\boldsymbol{y}_1 \equiv \Lambda^{-1/2}\Phi^T \boldsymbol{y} \tag{B.13}$$

generates two new random vectors with correlation matrices $R_x^1 = I, R_y^1$, respectively. Now let Ψ be the eigenvector matrix diagonalizing R_y^1. Then the random vectors generated by the unitary transformation ($\Psi^T\Psi = I$)

$$\boldsymbol{x}_2 \equiv \Psi^T \boldsymbol{x}_1 \tag{B.14}$$

$$\boldsymbol{y}_2 \equiv \Psi^T \boldsymbol{y}_1 \tag{B.15}$$

have correlation matrices $R_x^2 = I, R_y^2 = D$, where D is the diagonal matrix with elements the eigenvalues of R_y^1. Thus, the linear transformation of the original vectors by the matrix

$$A^T = \Psi^T \Lambda^{-1/2}\Phi^T \tag{B.16}$$

diagonalizes both correlation matrices simultaneously (one to an identity matrix). All these are obviously valid for covariance matrices as well. $\square$