MATH 271.1: Statistical Methods
Supplementary Notes on Principal Components Analysis

Jakov Ivan S. Dumbrique

1 Linear Algebra Review

This section presents the concepts and theorems from Linear Algebra which we will use in this class.

1. Eigenvectors and Eigenvalues
   Let $\mathbf{A}$ be an $n \times n$ matrix, and $\mathbf{x} \in \mathbb{R}^n$ a nonzero vector. If there exists a scalar value λ such that
   $$\mathbf{A}\mathbf{x} = \lambda \mathbf{x},$$
   then we define λ as an *eigenvalue* of $\mathbf{A}$, and the corresponding nonzero vector $\mathbf{x}$ is called an *eigenvector* of $\mathbf{A}$.

   To determine eigenvalues and eigenvectors, we solve the characteristic equation
   $$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$
   The eigenvectors of $\mathbf{A}$ corresponding to λ are the nonzero solutions of
   $$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$$

2. If c is a nonzero scalar and λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$ where $\mathbf{x}$ is its corresponding eigenvector, then $c\lambda$ is an eigenvalue of the scaled matrix $c\mathbf{A}$, and $\mathbf{x}$ is the eigenvector corresponding to the eigenvalue $c\lambda$.

3. Eigenvalues of Triangular Matrices
   If $\mathbf{A}$ is an $n \times n$ triangular matrix, then its eigenvalues are the entries on its main diagonal.

4. Diagonalizable Matrix
   An $n \times n$ matrix $\mathbf{A}$ is diagonalizable when $\mathbf{A}$ is similar to a diagonal matrix.
   That is, $\mathbf{A}$ is diagonalizable when there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P^{-1}AP}$ is a diagonal matrix.

5. If $\mathbf{A}$ and $\mathbf{B}$ are similar $n \times n$ matrices, then they have the same eigenvalues.

6. Condition for Diagonalization
   An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if it has n linearly independent eigenvectors.

7. Properties of Symmetric Matrices
   If $\mathbf{A}$ is an $n \times n$ symmetric matrix, then
   (a) $\mathbf{A}$ is diagonalizable
   (b) All eigenvalues of $\mathbf{A}$ are real

(c) If λ is an eigenvalue of $\mathbf{A}$ with multiplicity k, then λ has k linearly independent eigenvectors

(d) If λ_1 and λ_2 are distinct eigenvalues of $\mathbf{A}$, then their corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal.

8. Orthogonal Matrix
A square matrix $\mathbf{P}$ is orthogonal when it is invertible and $\mathbf{P^{-1}} = \mathbf{P^T}$.
In general, an $n \times m$ matrix $\mathbf{Q}$ is orthogonal if $\mathbf{Q^T Q} = \mathbf{I}$.

9. Orthogonally Diagonalizable Matrix
A matrix $\mathbf{A}$ is orthogonally diagonalizable when there exists an orthogonal matrix $\mathbf{P}$ such that $\mathbf{P^{-1}AP}$ is a diagonal matrix.

10. Fundamental Theorem of Symmetric Matrices
Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is orthogonally diagonalizable if and only if $\mathbf{A}$ is symmetric.

11. Diagonalizing a Square Matrix (Spectral Decomposition, Eigendecomposition)
If $\mathbf{A}$ is a diagonalizable square matrix, then $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \mathbf{PDP^{-1}}$$

where $\mathbf{P}$ is an $n \times n$ matrix whose columns consist of n linearly independent eigenvectors of $\mathbf{A}$, and $\mathbf{D}$ is a diagonal matrix with the corresponding eigenvalues of $\mathbf{A}$ on its main diagonal.

The order of the eigenvectors used to form $\mathbf{P}$ will determine the order in which the eigenvalues appear on the main diagonal of $\mathbf{D}$.

12. Positive Definite and Semidefinite Matrices

(a) A matrix $\mathbf{A}$ is *positive definite* if it is symmetric and $\mathbf{x^T A x} > 0$ for all nonzero $\mathbf{x}$.

(b) A matrix $\mathbf{A}$ is *positive semidefinite* if it is symmetric and $\mathbf{x^T A x} \geq 0$ for all $\mathbf{x}$.

13. Eigenvalues of Positive Definite and Semidefinite Matrices
The eigenvalues of a positive definite matrix are all positive. The eigenvalues of a positive semidefinite matrix are all nonnegative.

14. Derivatives of Matrices
If $\mathbf{u} = u(\mathbf{x})$, $\mathbf{v} = v(\mathbf{x})$, and $\mathbf{A}$ is not a function of $\mathbf{x}$, then the following properties hold true:

(a) $\dfrac{\partial\left(\mathbf{u} \cdot \mathbf{Av}\right)}{\partial \mathbf{x}} = \dfrac{\partial\left(\mathbf{u^T A v}\right)}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{Av} + \dfrac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{A^T u}$

(b) $\dfrac{\partial\left(\mathbf{u} \cdot \mathbf{v}\right)}{\partial \mathbf{x}} = \dfrac{\partial\left(\mathbf{u^T v}\right)}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{u}}{\partial \mathbf{x}}\mathbf{v} + \dfrac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{u}$