Robert Weismantel

Summary Lecture 27 and 28: Positive semidefinite matrices and the singular value decomposition

Preparations

The spectral theorem: Let A be a real $n \times n$ symmetric matrix

Let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors of A and $\lambda_1, \dots, \lambda_n$ the associated eigenvalues. Then $A = \sum_{i=1}^{n} \lambda_i v_i v_i^\top$

Proposition 9.2.10

Let $A \in \mathbb{R}^{n \times n}$ be symmetric. The Rayleigh Quotient, defined for $x \in \mathbb{R}^n \setminus \{0\}$, as

$$\text{For } x \in \mathbb{R}^n \setminus \{0\}, \text{ let } R(x) = \frac{x^\top A x}{x^\top x}.$$

R attains its maximum at $R(v_{\max}) = \lambda_{\max}$ and its minimum at $R(v_{\min}) = \lambda_{\min}$ where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of A and $v_{\max}$, $v_{\min}$ their associated eigenvectors.

Definition 9.2.11

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is said to be Positive Semidefinite / Positive Definite (PSD / PD) if all its eigenvalues are non-negative / positive.

Results about positive semidefinite matrices

Proposition 9.2.12

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is PSD if and only if $x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$.
A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is PD if and only if $x^\top Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Definition (Gram Matrix)

Given n vectors, $v_1, \ldots, v_n$ in $\mathbb{R}^m$, let $V \in \mathbb{R}^{m \times n}$ be the matrix with columns v_i. The Gram Matrix of V is the $n \times n$ matrix $G = V^\top V$.

Proposition 9.2.15

Let $A \in \mathbb{R}^{m \times n}$. The non-zero eigenvalues of $A^\top A \in \mathbb{R}^{n \times n}$ are the same as the ones of $AA^\top \in \mathbb{R}^{m \times m}$. Both matrices are also symmetric and PSD.

Proposition 9.2.16

Every symmetric positive semidefinite matrix M is a Gram matrix of an upper triangular matrix C. $M = C^\top C$ is known as the Cholesky Decomposition.

A sort of spectral theorem for general matrices?

Definition 9.3.1

Let $A \in \mathbb{R}^{m \times n}$. A singular value decomposition of A consists of orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that

$$A = U \Sigma V^\top, \tag{1}$$

where $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix, $U^\top U = I$ and $V^\top V = I$.
The columns of U (V) are the left (right) singular vectors of A. The diagonal elements of Σ, $\sigma_i = \Sigma_{ii}$ are called the singular values of A and are ordered as

$$\sigma_1 \geq \cdots \geq \sigma_{\min\{m,n\}} \geq 0.$$

Remark 9.3.2

If A has rank r we can write compactly $A = U_r \Sigma_r V_r^\top$, where $U_r \in \mathbb{R}^{m \times r}$ contains the first r left singular vectors, $V_r \in \mathbb{R}^{n \times r}$ contains the first r right singular vectors and $\Sigma_r \in \mathbb{R}^{r \times r}$ is diagonal with the first r singular values.

The SVD theorem?

The important idea

We will use the spectral theorem applied to the symmetric matrices $A^\top A$ and $AA^\top$. The singular values and vectors of A are in relation with eigenvalues and eigenvectors of these matrices.

Theorem (The SVD Theorem)

Every matrix $A \in \mathbb{R}^{m \times n}$ has an SVD decomposition of the form (1)*.*

In other words:

Every linear transformation is diagonal when viewed in the bases of the singular vectors.

Consequence of the SVD

Proposition 9.3.4

A rank-r matrix is a sum of r rank-1 matrices. Let $A \in \mathbb{R}^{m \times n}$ be a matrix of rank r. Let $\sigma_1, \dots, \sigma_r$ be the non-zero singular values of A with left and right vectors $u_1, \dots, u_r$, $v_1, \dots, v_r$, respectively. Then

$$A = \sum_{k=1}^{r} \sigma_k u_k v_k^\top . \qquad (2)$$

Final remarks

- The SVD is a powerful tool. Many results presented in this course become significantly simpler with the SVD.
- For instance, if A is invertible and A has SVD $A = U \Sigma V^\top$, then A^{-1} has SVD $A^{-1} = V \Sigma^{-1} U^\top$.
- Similarly, one can define the Moore-Penrose Pseudoinverse by using the SVD.