Linear Algebra
Week 13

G-07

19 XII 2024

1 Symmetric Matrices & Spectral Theorem

We diagonalized general matrices without any specific condition. The focus of this section will be symmetric matrices. Here I will list an overview of special attributes of symmetric matrices skipping the proofs which are included in the lecture notes.

Theorem 7.3.1. *(Spectral Theorem)* Any symmetric matrix $A \in \mathbb{R}^{n\times n}$ has n real eigenvalues and an orthonormal basis made of eigenvectors of A.

If you diagonalize a symmetric matrix along with Theorem 2.7.1. you get:

Corollary 7.3.2. For any symmetric matrix $A \in \mathbb{R}^{n\times n}$ there exists an orthogonal matrix $V \in R^{n\times n}$ (whose columns are eigenvectors of A) such that

$$A = V\Lambda V^\top$$

where $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix with the eigenvalues of A in its diagonal and $V^\top V = I$.

Using this eigendecomposition we can come to the conclusion that symmetric matrices can be written as sum of rank 1 matrices. In particular:

Proposition 7.3.6. Let A be a real $n \times n$ symmetric matrix and let $v_1, \ldots, v_n$ be an orthonormal basis of eigenvectors of A (the columns of the matrix V in Corollary 7.3.2.) and $\lambda_1, \ldots, \lambda_n$ the associated eigenvalues. Then

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^\top$$

In lecture notes you have the proof of the spectral theorem. Reading it is strongly recommended but I am not going to include it here again, it is the best that you read it from the original notes. There are some leading steps:

- **Proposition 7.3.7.** Let $A \in \mathbb{R}n \times n$ be a symmetric matrix and $\lambda \in \mathbb{C}$ an eigenvalue of A, then $\lambda \in \mathbb{R}$.
- **Corollary 7.3.8.** Every symmetric matrix $A \in \mathbb{R}n \times n$ has a real eigenvalue λ.
- **Remark 7.3.9.** proofs the fact that two eigenvectors of a symmetric matrix are orthogonal.
- The last step of the full proof is an induction proof and proves that "for any $1 \leq k \leq n$ there are k orthogonal eigenvectors of A".

In the following I listed some definitions regarding symmetric matrices from the lecture notes without proofs, which are as always recommended to read.

Rayleigh Quotient

Proposition 7.3.10. *(Rayleigh Quotient)* Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, the Rayleigh Quotient, defined for $x \in \mathbb{R}^n \setminus \{0\}$, as

$$R(x) = \frac{x^\top A x}{x^\top x}$$

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 respectively the largest and smallest eigenvalues of A, and $v_{\max}, v_{\min}$ their associated eigenvectors.

Be careful, we only define this for symmetric matrices as you can see in the definition. This is also the case for the following definition.

PD and PSD

Definition 7.3.11 (Positive Definite and Positive Semidefinite matrix). A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is said to be ***Positive Semidefinite (PSD)*** if all its eigenvalues are non-negative. If all the eigenvalues of A are strictly positive then we say A is ***Positive Definite (PD)***.

Using the Rayleigh Quotient we can derive a rule for positive (semi-)definiteness. Note that in Rayleigh Quotient the denominator can not be negative since $x^\top x = \|x\|^2$ and a square number is always non negative.

Proposition 7.3.12. A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is Positive Semidefinite if and only if $x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$. Analogously, a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is Positive Definite if and only if $x^\top Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Fact 7.3.13. If two $n \times n$ matrices A and B are PSD (or PD) then their sum is PSD (or PD).

Gram Matrix

Definition 7.3.14 (Gram Matrix). Given n vectors, $v_1, \ldots, v_n$ in $\mathbb{R}^m$, we call their Gram Matrix the $n \times n$ matrix of inner products $G_{ij} = v_i^\top v_j$.

Note that if $V \in \mathbb{R}^{m \times n}$ is the matrix whose columns are the n vectors, then $G = V^\top V$ is the Gram matrix of V.

Remark 7.3.15. Given a matrix $A \in \mathbb{R}^{m \times n}$, as an abuse of notation, we sometimes also call $AA^\top$ a Gram matrix of A. Notice that, if $a_1, \ldots, a_n \in \mathbb{R}^m$ are the columns of A, then $AA^\top$ is $m \times m$ and

$$AA^\top = \sum_{i=1}^{n} a_i a_i^\top.$$

The following proposition is an important result, especially in context of singular value decomposition.

Proposition 7.3.16. Given a real matrix $A \in \mathbb{R}^{m\times n}$, 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 symmetric and positive semidefinite.

The nonzero eigenvalues of $AA^\top$ and $A^\top A$ are the same. One of them has dimensions $m \times m$ and the other $n \times n$. So you can come to the conclusion that one of these matrices must have more eigenvalues than the other one if $m \neq n$. These additional eigenvalues are all 0.

Proposition 7.3.17. *(Cholesky decomposition)* 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.

2 Singular Value Decomposition

As you might have heard multiple times so far, this is "the ultimate theorem" of this course, singular value decomposition, or SVD. It is a powerful decomposition, that makes calculations and proofs easier in most cases. You are going to appreciate some of its true power in a 3^{rd} semester course, Numerical Methods.

In the lecture notes, SVD is introduced wonderful. Please read the whole section there, to fully understand the singular value decomposition. I will skip some crucial proofs and only give a glimpse of the theorem for the sake of completeness.

Definition 8.1.1 (SVD — Singular Value Decomposition). Let $A \in \mathbb{R}^{m\times n}$. There exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such that

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

where $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix, in the sense that $\Sigma_{ij} = 0$ when $i \neq j$, and the diagonal elements are non-negative and ordered in descending order. $U^T U = I$ and $V^T V = I$.
The columns $u_1, \ldots, u_m$ of U are called the *left singular vectors* of A and are orthonormal. The columns $v_1, \ldots, v_n$ of V are called the *right singular vectors* of A and are orthonormal. The diagonal elements of Σ, $\sigma_i = \Sigma_{ii}$, are called the *singular values* of A and are ordered as

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

- The left singular vectors of A, the columns of U, are the eigenvectors of $AA^\top$.
- The right singular vectors of A, the columns of V, are the eigenvectors of $A^\top A$.
- The nonzero singular values of A are square roots of the nonzero eigenvalues of $A^\top A$ or equivalently the nonzero singular values of A are equal to the square root of nonzero eigenvalues of $AA^\top$.
- $A \in \mathbb{R}^{m \times n}$ has $min\{m, n\}$ singular values including possibly repeating 0's.
- Number of non-zero singular values of a matrix is equal to its rank.

SVD tells us that we can write any rank r matrix as a sum of r rank-1 matrices. This takes away some of the computational work and storage issues if you look at it from a computer scientist's perspective.

Proposition 8.1.4. Let $A \in \mathbb{R}^{m \times n}$ be a matrix with rank r. Let $\sigma_1, \ldots, \sigma_r$ be the non-zero singular values of A, $u_1, \ldots, u_r$ the corresponding left singular vectors, and $v_1, \ldots, v_r$ the corresponding right singular vectors. Then

$$A = \sum_{k=1}^{r} \sigma_k u_k v_k^T.$$

As you might have realized, there are possibly many 0 rows or many 0 columns in the matrix Σ. For a rank r matrix you can store just the first r left singular vectors, first r left singular vectors and first r singular values. You have

$$A = U_r \Sigma_r V_r^\top$$

which stores less numbers and sometimes is called the *economical SVD*.

Theorem 8.1.5. (The SVD)
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 singular vectors.”

mkilic

♠