Lecture 9

Last time

- Missing Claim
- Concentration of the norm

Today

- Matrix norms
- Singular Value Decomposition
- Spectral Factorization

Matrix norms.

So far we have focused on scalars and vectors, in what follows we will deal with matrices. We need a linear algebra refresher.

We start with norms.

Def: Given a matrices $A, B \in \mathbb{R}^{n \times m}$, define the trace inner product as

$$\langle A, B \rangle = \mathrm{tr}(A^T B).$$

It induces the Frobenius norm

$$\|A\|_F^2 = \langle A, A \rangle.$$

This is equivalent to

$$\langle A, B \rangle = \mathrm{vec}(A)^T \mathrm{vec}(B) \quad \text{and} \quad \|A\|_F = \|\mathrm{vec} A\|_2.$$

We can also think of matrices as maps from $\mathbb{R}^m \to \mathbb{R}^n$ and this gives a different norm.

Def: The operator norm of A

$$\|A\|_{op} = \sup_{v \in S^{m-1}} \|Av\|_2.$$

Fact 0: Both $\|\cdot\|_F$ and $\|\cdot\|_{op}$ invariant under rotations, i.e. $\|UAV\| = \|A\|$.

$U \in O(n)$ $V \in O(m)$.

Matrix factorizations

One of the best ways to deal and think about matrices is via factorizations or decompositions.

Theorem (Singular Value Decomposition)

Let $A \in \mathbb{R}^{n \times m}$. Then, there exist orthogonal matrices $U \in O(n)$, $V \in O(m)$, and a diagonal matrix $\Sigma \in \mathbb{R}^{n \times m}$ such that

$$A = U \Sigma V^T$$

and the only (potentially) nonzero elements of Σ are

$$\Sigma_{11} \geq \Sigma_{22} \geq \cdots \geq \Sigma_{p} \geq 0$$

with $p = \min\{n, m\}$. $\dashv$

Remarks: One way to think of U and V^T is as change of coordinates that make

$$x \longmapsto Ax$$

look like a scaling of the axes.

Proof: Consider

$$(u_1, v_1) = \underset{\substack{u \in S^{n-1} \\ v \in S^{m-1}}}{\operatorname{argmax}} \; u^T A v$$

These two exist since $u^T A v$ is continuous and $S^{n-1} \times S^{m-1}$ is compact. Further, let $\sigma_1 = u_1^T A v_1$.

Claim: $Av_1 = \sigma_1 u_1$. (Why?)

From linear algebra we know we can construct orthonormal bases

$u_1, \tilde{u}_2, \dots, \tilde{u}_n$ of $\mathbb{R}^n$

$v_1, \tilde{v}_2, \dots, \tilde{v}_m$ of $\mathbb{R}^m$.

Let $\tilde{U} = \begin{pmatrix} | & | & & | \\ u_1 & \tilde{u}_2 & \dots & \tilde{u}_n \\ | & | & & | \end{pmatrix}$, $V = \begin{pmatrix} | & & | \\ v_1 & \cdots & \tilde{v}_m \\ | & & | \end{pmatrix}$.

Then,

$$\tilde{U}^T A \tilde{V} = \begin{bmatrix} \sigma_1 & w^T \\ 0 & \\ \vdots & B \\ 0 & \end{bmatrix} =: A_1.$$

$$(U^T A V)_{i1} = \langle \tilde{u}_i, A v_1 \rangle = \sigma \langle \tilde{u}_i, u_1 \rangle$$

Moreover,

$$\sigma_1 = \max_{\substack{u \in S^{n-1} \\ v \in S^{m-1}}} u^T A_1 v$$

(Why?)

$$= \max_{v \in S^{m-1}} \| A_1 v \|_2$$

$$\geq \left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \Big/ \| (\sigma_1, w) \|_2$$

$$\geq \sqrt{\sigma_1^2 + \|w\|_2^2}$$

$$\geq \sigma_1 .$$

Thus, we conclude $w = 0$. Hence, $\tilde{U}^T A \tilde{V} = \begin{bmatrix} \sigma & 0 \\ 0 & B \end{bmatrix}$. Repeating the argument inductively, we get $\exists\ U \in O(n)$, $V \in O(m)$

$$U^T A V = \Sigma \qquad \text{(Check!)}$$

with Σ diagonal as we wanted. Then, it is immediate that $A = U \Sigma V^T$. $\square$

The vectors

$$U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} \text{ and } V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_m \\ | & & | \end{pmatrix}$$

are called left and right-singular vectors respectively.

Further $\sigma_i = \Sigma_{ii}$ is the ith singular value of A. We use $\sigma_{max}(A)$, $\sigma_{min}(A)$ for the maximum and minimum singular values, respectively.

Corollary (Properties)

Let $A \in \mathbb{R}^{n \times m}$ and let $U \Sigma V^T$ be its SVD decomposition. Then,

1) $Av_i = \sigma_i u_i$ and $A^T u_i = \sigma_i v_i$.

2) For any $x \in \mathbb{R}^n$,
$$\sigma_{min}(A) \|x\|_2 \leq \|Ax\|_2 \leq \sigma_{max}(A) \|x\|_2.$$

3) Let $r = \operatorname{rank} A$, then
$$\sigma_r(A) > \sigma_{r+1}(A) = 0.$$

4) We have
$$\operatorname{null} A = \operatorname{span}\{v_{r+1}, \ldots, v_m\},$$
$$\operatorname{range} A = \operatorname{span}\{u_1, \ldots, u_r\}.$$

5) We can write
$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T.$$

6) $\|A\|_F = \|\sigma\|_2$ and $\|A\|_{op} = \|\sigma\|_\infty$.

$\sigma = \mathrm{vec}(\sigma_1, \sigma_2, \dots)$ $\qquad$ $\|\sigma\|_\infty = \max_i |\sigma_i|$. $\dashv$

For symmetric matrices

$$\mathbb{S}^n := \{A \in \mathbb{R}^{n\times n} \mid A = A^T\}$$

we have another useful factorization

Theorem (Spectral decomposition): Let $A \in \mathbb{S}^n$. Then, there exists $U \in O(n)$ and $\Sigma = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{S}^n$ s.t.

$$A = U \Lambda U^T.$$

$\dashv$

Exercise: Prove that symmetric matrices have real eigenvalues and their eigenspaces are orthogonal to each other. Use this fact to prove the theorem above. $\dashv$