1 Overview

In the last lecture an outline and motivation for the course was given. In this lecture we review Singular Value Decomposition (SVD) and its perturbation bounds. Additionally, we introduce the Semi-definite Programming (SDP) and the topic of Convexity.

2 Singular Value Decomposition (SVD)

Theorem 1. Let $A \in \mathbb{C}^{m \times N}$. Then $A$ has a “unique” SVD:

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

where

i) $U \in \mathbb{C}^{m \times m}$ is unitary, i.e. $U^{-1} = U^*$

ii) $V \in \mathbb{C}^{N \times N}$ is unitary
iii) \( \Sigma \in \mathbb{R}^{m \times N} \) is diagonal, i.e.

\[
\Sigma = \begin{pmatrix}
\sigma_1 & 0 & \cdots & 0 \\
0 & \sigma_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sigma_q
\end{pmatrix}
\]

where \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q \) and \( q = \min(m, N) \).

The “unique” diagonal entries \( \sigma_1, \sigma_2, \ldots, \sigma_q \) are called the singular values of \( A \).

Proof: The proof follows from the next 2 lemmas.

Before presenting the 2 lemmas, we begin by some definitions.

**Definition 1.** Let \( \{ w_1, \cdots, w_N \} \) be an orthonormal basis for \( \mathbb{C}^N \). Define

\[
h_j = \begin{cases}
0 & \text{if } s_j = 0 \\
\frac{1}{s_j} A w_j & \text{if } s_j \neq 0
\end{cases}
\]

and define \( W \) to be the unitary matrix \( (w_1 \cdots w_N) \in \mathbb{C}^{N \times N} \).

**Lemma 1.**

\[
A = (h_1 \cdots h_N) \begin{pmatrix}
\sigma_1 & 0 & \cdots & 0 \\
0 & \sigma_2 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \sigma_q & 0
\end{pmatrix} W^*
\]

Proof:

\[
AW = \begin{pmatrix}
\sigma_1 & 0 & \cdots & 0 \\
0 & \sigma_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sigma_q
\end{pmatrix} (h_1 \cdots h_N)
\]

and \( W^{-1} = W^* \).

**Lemma 2.** Let \( A \in \mathbb{C}^{m \times N} \). Let \( w_1, \cdots, w_N \) be the eigenvectors of \( A^* A \) (which is a Hermitian). Then \( \langle h_j, h_l \rangle = 0 \) if \( j \neq l \).

Proof:

\[
\langle Aw_j, Aw_l \rangle = (Aw_j)^* Aw_l \\
= w_j^* A^* A w_l \\
= w_j^* (\lambda_l w_l) \\
= 0 \text{ if } j \neq l.
\]
2.1 Calculation of the SVD

Notice that

\[ A^*A = (U \Sigma V^*)^*(U \Sigma V^*) \]
\[ = V \Sigma^2 V^* \]

Then, \( V \) contains the eigenvectors of \( A^*A \) as columns, and \( \sigma_1, \sigma_2, \ldots, \sigma_q \) are the squared eigenvalues of \( A^*A \).

Numerically, we can use, e.g., the QR algorithm to find the eigenvalues of \( A^*A \) to get the singular values of \( A \). The shifted inverse power method, e.g., can be used to calculate \( V \). Similarly, from \( AA^* \) we can find \( U \).

3 Perturbation Bounds

**Theorem 2. Weyl’s Bounds**

Let \( A, E \in \mathbb{C}^{m \times N} \) and \( q = \min(m, N) \). Then the following inequalities hold (Notice that the singular values are assumed to be ordered)

a) \( \sigma_{j+i-1}(A + E) \leq \sigma_i(A) + \sigma_j(E) \)

b) \( \sigma_{j+i-1}(AE^*) \leq \sigma_i(A)\sigma_j(E) \)

for \( 1 \leq i, j \leq q \) such that \( i + j \leq q + 1 \).

**Proof:** For a proof, see [HJ94].

There are perturbation results for the singular vectors as well. (See Stewart’s notes.)

3.1 Homework Problems

**Problem 1** Choose scribe dates.

**Problem 2** Prove using Theorem 2 that

\[ |\sigma_i(A + E) - \sigma_i(A)| \leq \sigma_1(E). \]

**Problem 3** Suppose every entry of \( A \in \mathbb{C}^{m \times N} \) is corrupted with an additive error of magnitude \( \leq \varepsilon \). How much is \( \sigma_1(A) \) changed in terms of \( \varepsilon \)?

4 Semi-definite Programming (SDP)

**Standard Form** (See [VB’96])

Given \( c \in \mathbb{R}^m \) and \( F_0, \ldots, F_m \in \mathbb{S}^N \) that are fixed. (Note: \( \mathbb{S}^N \) is the space of real symmetric \( N \times N \) matrices.)
And given variables $x \in \mathbb{R}^m$. Minimize $c^T x$ such that

$$F(x) = F_0 + \sum_{j=1}^{m} x_j F_j \geq 0$$

which means the matrix $F(x)$ is positive semi-definite.

Notes

- $F(x) \in \mathbb{S}^N$ and so $F(x)$ has $N$ eigenvalues (with possible repetitions).
- $F(x) \geq 0$ tells us that we want all the $N$ eigenvalues to remain non-negative (within machine precision).
- $y^T F(x) y \geq 0 \ \forall y \in \mathbb{R}^N$.
- SDPs can be solved computationally in polynomial time in $mN$. (See “Interior Point Methods”.)
- SDPs are a subset of the more general convex optimization problems.

5 Convexity

(See [BV04] and Appendix B of [FR13]).

5.1 Definitions

Definition 2. A function $f : \mathbb{R}^D \to \mathbb{R}^d$ is convex if

$$f(\alpha x + \beta y) \leq (\star) \alpha f(x) + \beta f(y)$$

$\forall x, y \in \mathbb{R}^D \text{ and } \forall \alpha, \beta \in (0, 1)$ such that $\alpha + \beta = 1$.

$(\star)$ The inequality is coordinate-wise.

- $f(x) = c^T x$ is convex since it is linear.

Definition 3. A function $F : \mathbb{R}^m \to \mathbb{R}^{N \times N}$ is convex if

$$F(\alpha x + \beta y) \leq (\star) \alpha F(x) + \beta F(y)$$

$\forall x, y \in \mathbb{R}^m \text{ and } \forall \alpha, \beta \in \mathbb{R}^+$ such that $\alpha + \beta = 1$.

$(\star)$ For matrices, $A \leq B$ if $B - A$ is positive semi-definite.

- $F(x) = F_0 + \sum_{j=1}^{m} x_j F_j$ is convex since it is linear.
Definition 4. A set $K \subseteq \mathbb{R}^m$ is convex if

$$(\alpha x + (1 - \alpha)y) \in K$$

$\forall x, y \in K$ and $\forall \alpha \in (0, 1)$.

- The set $K = \{x| F(x) \geq 0\}$ used in the SDP constraint is a convex set.

Definition 5. $K \subseteq \mathbb{R}^m$ is a convex cone if it is both convex and a cone.

Definition 6. $K \subseteq \mathbb{R}^m$ is a cone if $\alpha x \in K$, $\forall x \in K$ and $\forall \alpha \in \mathbb{R}^+$.

5.2 Homework Problems

Problem 4 Show that $K = \{x| F(x) = F_0 + \sum_{j=1}^m x_j F_j\}$ (used in the SDP constraint) is convex.

References

