Matrix Analysis: Singular value decomposition and applications

Jean-Luc Bouchot

School of Mathematics and Statistics
Beijing Institute of Technology
jlbouchot@bit.edu.cn

2018/11/18

Outline

Definition 2.1

A complex quadratic form is said to be **positive-definite** *if* $f(\mathbf{x}) = \mathbf{x}^* A\mathbf{x} > 0$*, for all* $\mathbf{x} \in \mathbb{K}^n \backslash \{0\}$*. On this case, the (Hermitian) matrix* A *is said to be* **positive definite***.*
If the strict inequality is relaxed to ≥ 0*, we say that the quadratic form* f *and the matrix* A *are* **positive semi definite***.*

Definition 2.2

If the (Hermitian) matrix $-A$ is positive (semi-)definite, then A is said to be **negative (semi-)definite**.

Theorem 2.1

Let A be a Hermitian matrix and let f denote its associated complex quadratic form. The following statements are equivalent.

1. *f is positive definite.*
2. *C^*AC is positive definite for every invertible C.*
3. $\sigma(A) \subset \mathbb{R}_{\geq 0}$.
4. *There exists an invertible matrix P such that $P^*AP = I$.*
5. *There exists an invertible matrix Q such that $A = Q^*Q$.*

Exercise 2.1

Show that if A is a unitary positive definite matrix, then $A = I$.

Exercise 2.2

Let A and B be two positive semi-definite matrices. Prove twice that $A + B$ is positive semi definite using

1. a direct computation of the associated complex quadratic form,
2. Weyl's inequalities

Theorem 2.2

Let A and B be two Hermitian matrices and assume moreover that B is positive definite. Then there exists a non-singular matrix P such that

$$P^* A P = \operatorname{diag}(\alpha_1, \cdots, \alpha_n), \quad P^* B P = I.$$

The scalars $\alpha_1, \cdots, \alpha_n$ are independent of the matrix P.

Lemma 1

Let $A \in \mathbb{K}^{m\times n}$. *It holds*

$$rk(A) = rk(A^*) = rk(AA^*) = rk(A^*A).$$

Remark 2.1

In case of real matrices, we can replace the Hermitian conjugate by simple transpose. Verify that this cannot be true for complex matrices.

Lemma 2

Let A be an $m \times n$ matrix and B be an $n \times m$ matrix and assume $m \leq n$. The following holds

$$p_{BA}(t) = t^{n-m} p_{AB}(t).$$

Corollary 2.1

Let $A \in \mathbb{K}^{m \times n}$. *Then* AA^* *and* A^*A *have the same spectrum (up to the* 0 *eigenvalue and its multiplicity).*

Lemma 3

AA^* and A^*A are positive semidefinite matrices.

Definition 2.3

The singular values of a matrix $A \in \mathbb{K}^{m\times n}$ *are the square roots of the eigenvalues of* AA^* *or equivalently* A^*A*:*

$$\sigma_i(A) = \sqrt{\lambda_i(AA^*)} = \sqrt{\lambda_i(A^*A)}, 1 \leq i \leq \min\{m, n\}.$$

Remark 2.2

It is important to notice the followings:

1. assuming the eigenvalues to be enumerated in non-decreasing order, we have that $\lambda_i(AA^*) = \lambda_i(A^*A) \geq 0$.
2. the number of non-zero singular values are precisely the rank of A: r.
3. the number of (counting multiplicities) singular values is equal to the smallest dimension – this explains the somewhat odd definition
4. only the non-zero singular eigenvalues will often play a role. The singular value decomposition below will make it clear what is meant.

Exercise 2.3

Singular values and eigenvalues usually not related. Find the eigenvalues and singular values of the following matrices:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Lemma 4

Let $A \in \mathbb{K}^{m\times n}$*. There exists an orthonormal basis* $\mathbf{u}_1, \cdots, \mathbf{u}_n$ *of* $\mathbb{K}^n$ *such that the family* $\{A\mathbf{u}_1, \cdots, A\mathbf{u}_n\}$ *is orthogonal (in* $\mathbb{K}^m$*).*

Lemma 5

Let $A \in \mathbb{K}^{m \times n}$ *and let* $(\mathbf{u}_1, \cdots, \mathbf{u}_n)$ *be an orthonormal basis of* $\mathbb{K}^n$*. Define*

$$\mathbf{v}_j = \begin{cases} \frac{A\mathbf{u}_i}{\|A\mathbf{u}_i\|} & \textit{if} \quad \|A\mathbf{u}_i\| \neq 0 \\ 0 & \textit{if} \quad \|A\mathbf{u}_i\| = 0 \end{cases}$$

Let $D = \mathrm{diag}(\|A\mathbf{u}_1\|, \cdots, \|A\mathbf{u}_n\|)$*. Define* $U = [\mathbf{u}_1|\cdots|\mathbf{u}_n]$ *and* $V = [\mathbf{v}_1|\cdots|\mathbf{v}_n]$*. The matrix* A *enjoys the decomposition*

$$A = VDU^*.$$

Lemma 6

Let $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ *be an orthormal basis of eigenvectors for* A^*A*. Then* $\|A\mathbf{u}_i\| = \sqrt{\lambda_i}$.

Definition 2.4

Let $A \in \mathbb{K}^{m \times n}$ *be a matrix. There exists a* **Singular Value Decomposition (SVD)**

$$A = VSU^*$$

such that $V \in \mathbb{K}^{n \times n}$ *is an orthonormal basis of eigenvector of* AA^**,*
$U \in \mathbb{K}^{m \times m}$ *is an orthonormal basis of eigenvectors of* A^*A *and* $S \in \mathbb{K}^{m \times n}$ *is a matrix such that* $d_{ii} = \sqrt{\lambda_i}$*, for* $1 \leq i \leq \min\{m, n\}$ *and* 0 *elsewhere.*
U *and* V *are respectively called the right and left singular vectors.*

Exercise 2.4

Prove that $A = \sum_{i=1}^{\min\{m,n\}} \sqrt{\lambda_i}\mathbf{v}_i\mathbf{u}_i^*$.

Proposition 2.1 (Truncated SVD)

Let $A \in \mathbb{K}^{m \times n}$ *be a rank* r *matrix. There exists* r *strictly positive numbers* $\sigma_1, \cdots, \sigma_r$*,* r *orthonormal vectors in* $\mathbb{K}^m$ $\mathbf{v}_1, \cdots, \mathbf{v}_r$ *(column wise in a matrix* V_r*) and* r *orthonormal vectors* $\mathbf{u}_1, \cdots, \mathbf{u}_r$ *(stacked in* U_r*) such that*

$$A = V_r S_r U_r^*$$

where $S_r = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$*.*

Definition 2.5

Let $A \in \mathbb{K}^{m \times n}$ be a matrix and let $A = VSU^*$ be its singular value decomposition in which the singular values are numbered in decreasing order of magnitude. $A_k = V_k S_k U_k^*$ is called the **(best) rank k approximation of A**.

Exercise 2.5

Find the singular value decompositions and the rank 1 and 2 approximations of the following matrices.

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix}.$$

Definition 2.6

A norm $\|\cdot\|$ on the vector space of matrix is said to be a **matrix norm** *(by opposition to a vector norm) if*

1. *it is a vector norm*
2. *it is submultiplicative: for any two matrices A and B such that the product AB is defined, $\|AB\| \leq \|A\|\|B\|$.*

Proposition 2.2

The Frobenius norm is a matrix norm.

Exercise 2.6

Show that the following are matrix norms:

1. $\|A\| = \sum_{i,j} |a_{i,j}|$.
2. $A \in \mathbb{K}^{n\times n}$, $\|A\| = n \max i, j |a_{i,j}|$.

Definition 2.7

Let $\| \cdot \|_V$ be a vector norm and $\| \cdot \|_M$ be a matrix norm. $\| \cdot \|_V$ is **compatible** *with $\| \cdot \|_M$ if*

$$\|A\mathbf{x}\|_V \leq \|A\|_M \|\mathbf{x}\|_V.$$

Exercise 2.7

Verify that the $\|\cdot\|_2$ norm is compatible with the Frobenius norm.

Definition 2.8

Let $\|\cdot\|_{k\to k}$ *be defined on the set of matrices as*

$$\|A\|_{k\to k} := \max_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_k}{\|\mathbf{x}\|_k}.$$

$\|\cdot\|_{k\to k}$ *is called the* **operator norm induced by** $\|\cdot\|_V$.

Theorem 2.3

The operator norm is indeed a matrix norm and the vector norm used is compatible with it.

Remark 2.3

1. The operator norms need not consider the same norm in the input and output spaces ... but in this case, we need to review the definition of the compatibility, which is not important enough here.
2. The k norm can be pretty much any thing, and not necessarily the Euclidean or other Minkowski norms.

Proposition 2.3

Operator norms induced by some Minkowski norms are quite common and should be understood:

1. *The maximum column sum norm is induced by the* 1 *norm:* $\|A\|_{1\to 1} = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$.
2. *The spectral norm is* $\|A\_2 \to 2 = \max_j \sqrt{\lambda_j(A^*A)}$ *is the largest singular value.*
3. *The maximum row sum is induced by the infinity norm:* $\|A\|_{\infty\to\infty} = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{i,j}|$.

Lemma 7

Let B be a rank k matrix. Then

$$\|A - A_k\|_F \leq \|A - B\|_F$$

Lemma 8

$$\|A - A_k\|_{2\to 2}^2 = \sigma_{k+1}^2.$$

Theorem 2.4

Let $A \in \mathbb{K}^{m \times n}$ and let B be a rank k matrix. Then

$$\|A - A_k\|_{2\to 2} \leq \|A - B\|_{2\to 2}.$$

Said differently, the truncated matrix A_k is the best rank k approximation of A when measured in the $2 \to 2$ norm.

Definition 2.9

Let $A \in \mathbb{K}^{m\times n}$. A matrix $B \in \mathbb{K}^{n\times m}$ is said to be a **pseudo-inverse** *if it satisfies the following axioms:*

1. $ABA = A$,
2. $BAB = B$,
3. BA *and* AB *are Hermitian.*

Theorem 2.5

Let $A \in \mathbb{K}^{m \times n}$. If there exists such a pseudo inverse, it is unique.

Theorem 2.6

Let $A \in \mathbb{K}^{m \times n}$ *be a matrix. There always exists a pseudo inverse.*

Theorem 2.7

Assume A is a square non-singular matrix. Then $A^\dagger = A^{-1}$.

Proposition 2.4

Let $A \in \mathbb{K}^{m \times n}$

1. *(overdetermined systems, more equations than unknown) If* $m \geq n$ *and* A *has full rank* (n)*, then* $A^\dagger = (A^*A)^{-1}A^*$*. It follows* $A^\dagger A = I_n$*.*
2. *(underdetermined systems) If* $m \leq n$ *and* A *has full rank* (m)*, then* $A^\dagger = A^*(AA^*)^{-1}$*. It follows* $AA^\dagger = I_m$*.*