

Lecture 4

SVD I

Last lecture we computed a positive scalar, σ_1 , that characterizes the magnitude of maximum amplification that a matrix A can effect on any vector. The singular value decomposition does this and more:

singular values are the proportions of maximum amplification of orthogonal directions:

1. σ_1 is a maximum amplification of any direction.
2. σ_2 is a maximum amplification of any direction orthogonal to the direction σ_1 amplifies.
3. σ_3 is a maximum amplification of any direction orthogonal to the direction σ_1 amplifies *and* the direction σ_2 amplifies.
4. ...
5. $\sigma_{\text{rank } A}$ is a minimum amplification of any direction not in the null space of A .

right singular vectors are the directions that are detected for amplification: v_1 is how σ_1 finds what it wants, etc.

left singular vectors are the directions that are output after amplification: σ_1 detects the quantity of v_1 in its input, amplifies it, and outputs its amplified version along u_1 , etc.

The SVD can be seen as a generalization of the maxim “a vector is magnitude and direction” to matrices: “a matrix is magnitudes, input directions, and output directions.” It is customary to state the magnitudes and directions of a matrix in order of importance.

Definition 19 (SVD, abstract version). *Let $f : V \rightarrow U$ be a linear map of inner product spaces. A **singular value decomposition** of f is a choice of*

- an orthonormal basis $\{v_1, v_2, \dots, v_m\}$ for V (**right singular vectors**)
- an orthonormal basis $\{u_1, u_2, \dots, u_n\}$ for U (**left singular vectors**), and
- positive scalars $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ (**singular values**); such that

$$f(v_i) = \sigma_i u_i \text{ for } 1 \leq i \leq r = \text{rank } f.$$

Lemma 5 (A “square”¹ of a map). *Let $f : V \rightarrow U$ be a linear map of inner product spaces.*

1. The composite $f^* f : V \rightarrow V$ is self-adjoint.
2. The eigenvalues of $f^* f$ are nonnegative.
3. The null space of $f^* f$ is the same as that of f .

¹In the same sense that $\bar{z}z$ is a nonnegative real number whose “size” is the square of the complex number z .

4. The rank of f^*f is the same as that of f .

Proof. 1. Proved yesterday.

2. Let $f^*f v = \lambda v$. Then $\langle f^*f v, v \rangle = \langle f v, f v \rangle = \lambda \langle v, v \rangle$, establishing λ as a ratio of nonnegative numbers.

3. Obviously the null space of f is contained in the null space of f^*f . We'll show that the null space of f^*f is contained the null space of f . Let $f^*f v = 0$. Then $\langle f^*f v, v \rangle = \langle f v, f v \rangle = 0$ and $f v = 0$.

4. Follows from previous part by the rank-nullity theorem. \square

Theorem 16 (SVD exists). Let $f : V \rightarrow U$ be a linear map of vector spaces, with $\dim V = n$ and $\dim U = m$.

Proof. (Construction)

1. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ be the positive eigenvalues of f^*f , and let $\{v_1, v_2, \dots, v_r\}$ be corresponding orthonormal eigenvectors.
2. Choose $\sigma_i = \sqrt{\lambda_i}$ for $i = 1 \leq i \leq r$.
3. Choose $u_i = f(v_i)/\sigma_i$ for $i = 1 \leq i \leq r$.
4. Complete $\{v_i\}_{1 \leq i \leq r}$ and $\{u_i\}_{1 \leq i \leq r}$ to orthonormal bases for V and U , respectively (the former having n vectors and the latter m).

(Verification) The basis $\{v_i\}$ for V is orthonormal, the purported singular values are in nonincreasing order, and $f(v_i) = \sigma_i u_i$ by construction.

All that remains to show is that $\{u_i\}_{1 \leq i \leq m}$ is orthonormal. By construction, $\{u_i\}_{n < i \leq m}$ is orthonormal, and is orthogonal to $\{u_i\}_{1 \leq i \leq r}$. So we just have to show that $\{u_i\}_{1 \leq i \leq r}$ is orthonormal. For $1 \leq i, j \leq r$,

$$\langle u_i, u_j \rangle = \left\langle \frac{f(v_i)}{\sigma_i}, \frac{f(v_j)}{\sigma_j} \right\rangle \quad (4.1)$$

$$= \frac{1}{\sigma_i \sigma_j} \langle (f^*f)(v_i), v_j \rangle \quad (4.2)$$

$$= \frac{1}{\sigma_i \sigma_j} \langle \sigma_i^2 v_i, v_j \rangle \quad (4.3)$$

$$= \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (4.4)$$

\square

Example: SVDs of the identity map/matrix

Let's find out what the SVDs of $I \in \mathbb{C}^{n \times n}$ are.

1. We have to form the matrix I^*I . It equals I .
2. We need a basis of orthonormal eigenvectors for $I^*I = I$. Any basis $\{v_1, v_2, \dots, v_n\}$ will do. (One could make such a basis using the Gram-Schmidt process by pulling random vectors out of hat.) Every eigenvalue of I^* is 1, so every singular value is 1.
3. The left singular vectors are I times the right singular vectors, divided by 1.

Therefore, the singular value decompositions of I are all orthonormal bases.

SVD as a matrix factorization

Let $A \in \mathbb{C}^{m \times n}$. Define $V \in \mathbb{C}^{n \times n}$ by

$$V = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}, \quad (4.5)$$

where v_1, v_2, \dots, v_n are right singular vectors in an SVD of A . Define

$$U = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix}, \quad (4.6)$$

where u_1, u_2, \dots, u_m are right singular vectors such that $f(v_i) = \sigma u_i$. Define a matrix $\Sigma \in \mathbb{C}^{m \times n}$ by

$$\Sigma_{ii} = \sigma_i. \quad (4.7)$$

Then the following factorization holds:

$$A = U \Sigma V^*. \quad (4.8)$$

Splitting this matrix product into its nonzero outer products,

$$A = \sum_{i=1}^{\text{rank } A} \sigma_i (u_i v_i^*). \quad (4.9)$$

This form is powerful because it allows us to approximate A by simpler matrices. Suppose that A is a noisy measurement of a matrix that we earnestly believe to be of rank $r' < \text{rank } A$. (Full rank matrices are the “hay in the haystack,” so noise tends to be full rank.)

$$A \stackrel{\text{low rank}}{\approx} \sum_{i=1}^{r'} \sigma_i (u_i v_i^*). \quad (4.10)$$

Or suppose that A is a noisy measurement of a matrix we believe in truth to be unitary. (Unitary matrices are common in robotics.) Probably all of the singular values are close to 1. To get the nearest unitary matrix, we can set them all to 1.

$$A \stackrel{\text{unitary}}{\approx} \sum_{i=1}^{\text{rank } A} u_i v_i^*. \quad (4.11)$$