MMAT 5320 Computational Mathematics

Eric Chung

Lecture 1

Let x be an n-dimensional column vector, A be an $m \times n$ matrix.

The vector $b = Ax$ is the m-dimensional column vector defined as

$$b_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \cdots, m$$

where b_i is the i-th entry of the vector b.

We assume the entries are complex numbers, denoted $\mathbb{C}$.

The space of m-vectors is $\mathbb{C}^m$, and the space of $m \times n$ matrices is $\mathbb{C}^{m \times n}$.

The map $x \mapsto Ax$ is linear, meaning that, for any $x, y \in \mathbb{C}^n$ and any $\alpha \in \mathbb{C}$,

$$\begin{aligned} A(x+y) &= Ax + Ay \\ A(\alpha x) &= \alpha Ax \end{aligned}$$

Conversely, any linear map from $\mathbb{C}^n$ to $\mathbb{C}^m$ can be expressed by a $m \times n$ matrix.

Let a_j be the j-th column of the matrix A. Then $b = Ax$ is given by

$$b_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \cdots, m$$

which is the same as

$$b = \sum_{j=1}^{n} x_j a_j$$

$$\begin{bmatrix} \\ b \\ \\ \end{bmatrix} = \begin{bmatrix} \\ a_1 \,\Big|\, a_2 \,\Big|\, \cdots \,\Big|\, a_n \\ \\ \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} \\ a_1 \\ \\ \end{bmatrix} + x_2 \begin{bmatrix} \\ a_2 \\ \\ \end{bmatrix} + \cdots + x_n \begin{bmatrix} \\ a_n \\ \\ \end{bmatrix}$$

That is, b is a linear combination of the columns a_j.

For the matrix-matrix product $B = AC$, each column of B is a linear combination of the columns of A.

To see this, if A is $\ell \times m$ and C is $m \times n$, then B is $\ell \times n$ and

$$b_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj}$$

In terms of columns:

$$\left[\begin{array}{c|c|c|c} \boldsymbol{b_1} & \boldsymbol{b_2} & \cdots & \boldsymbol{b_n} \end{array}\right] = \left[\begin{array}{c|c|c|c} \boldsymbol{a_1} & \boldsymbol{a_2} & \cdots & \boldsymbol{a_m} \end{array}\right] \left[\begin{array}{c|c|c|c} \boldsymbol{c_1} & \boldsymbol{c_2} & \cdots & \boldsymbol{c_n} \end{array}\right]$$

The above can be written as

$$b_j = Ac_j = \sum_{k=1}^{m} c_{kj} a_k$$

Example 1.2: Outer product. Product of an m-dimensional column vector u with an n-dimensional row vector v.

$$\begin{bmatrix} \\ u \\ \\ \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \left[\begin{array}{c|c|c|c} & & & \\ v_1 u & v_2 u & \cdots & v_n u \\ & & & \end{array}\right] = \begin{bmatrix} v_1 u_1 & \cdots & v_n u_1 \\ \vdots & & \vdots \\ v_1 u_m & \cdots & v_n u_m \end{bmatrix}$$

The columns are multiples of u, the rows are multiples of v.

Example 1.3: Consider $B = AR$, where R is the $n \times n$ upper triangular matrix with entries $r_{ij} = 1$ for $i \leq j$ and $r_{ij} = 0$ for $i > j$. The product is illustrated as

$$\left[\begin{array}{c|c|c} & & \\ b_1 & \cdots & b_n \\ & & \end{array}\right] = \left[\begin{array}{c|c|c} & & \\ a_1 & \cdots & a_n \\ & & \end{array}\right] \begin{bmatrix} 1 & \cdots & 1 \\ & \ddots & \vdots \\ & & 1 \end{bmatrix}$$

The columns are

$$b_j = Ar_j = \sum_{k=1}^{j} a_k$$

The matrix R is a discrete analog of an indefinite integral operator.

Range and rank of a matrix

The range of a matrix A, denoted $\text{range}(A)$, is the set of vectors that can be expressed as Ax for some x.

Theorem: $\text{range}(A)$ is the space spanned by the columns of A.

Note:

- the range of a matrix A is also called the column space of A
- the nullspace of A, denoted $\text{null}(A)$, is the set of vectors x such that $Ax = 0$

The column rank of a matrix is the dimension of its column space. The row rank of a matrix is the dimension of the space spanned by its rows. They are always the same, so we refer to this number as the rank of a matrix.

Note:

- an $m \times n$ matrix has full rank if it has the maximal possible rank (the least of m and n)

Inverse

A nonsingular or invertible matrix is a square matrix of full rank.

Note that the m columns of a nonsingular $m \times m$ matrix A form a basis for the whole space $\mathbb{C}^m$. Thus, any vector can be uniquely expressed in terms of these columns.

Let e_j be the canonical unit vector, that is, its j-th entry equals one, and zeros elsewhere. We can write

$$e_j = Az_j$$

or

$$\left[\begin{array}{c|c|c} \\ e_1 & \cdots & e_m \\ \\ \end{array}\right] = I = AZ$$

Here, I is the identity matrix.

The matrix Z is called the inverse of A. Note that, any nonsingular matrix A has a unique inverse, denoted as A^{-1}. It satisfies $AA^{-1} = A^{-1}A = I$.

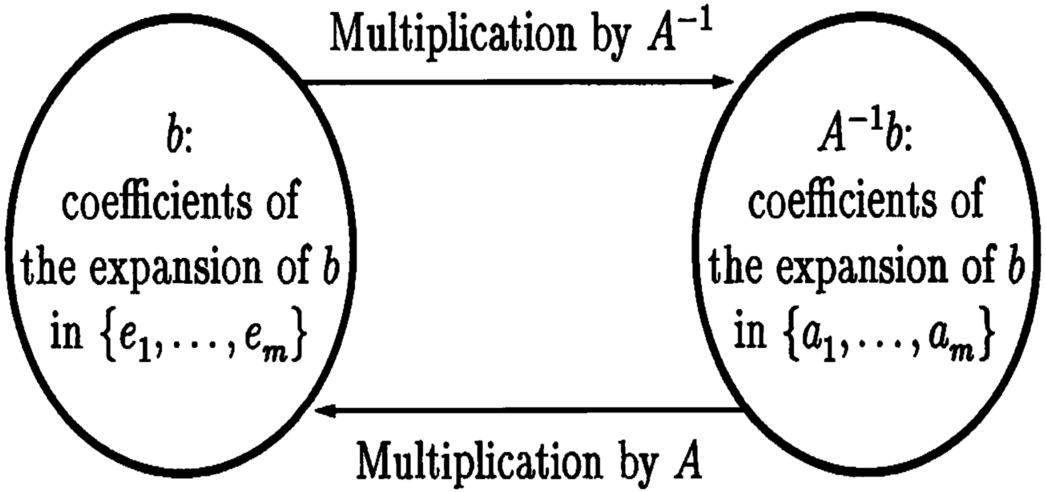
Some properties

Theorem: Let $A \in \mathbb{C}^{m\times m}$. The following are equivalent

1. A has an inverse A^{-1}
2. $\text{rank}(A) = m$
3. $\text{range}(A) = \mathbb{C}^m$
4. $\text{null}(A) = \{0\}$
5. 0 is not an eigenvalue of A
6. 0 is not a singular value of A
7. $\det(A) \neq 0$

An important observation:

Let $x = A^{-1}b$. Then we have $Ax = b$, that is, x is the vector of coefficients of the unique linear expansion of b in the basis of columns of A.



Lecture 2

The hermitian conjugate or adjoint of an $m \times n$ matrix A, denoted A^*, is the $n \times m$ matrix whose i, j entry is the complex conjugate of the j, i entry of A. For example,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \implies A^* = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{21} & \bar{a}_{31} \\ \bar{a}_{12} & \bar{a}_{22} & \bar{a}_{32} \end{bmatrix}$$

where $\bar{a}$ is the complex conjugate of a.

If $A = A^*$, then A is hermitian.

For real matrix A, the adjoint is also known as the transpose, denoted A^T. If a real matrix A is hermitian, then A is called symmetric.

Inner product

The inner product of two column vectors $x, y \in \mathbb{C}^m$ is

$$x^* y = \sum_{i=1}^{m} \overline{x}_i y_i$$

The Euclidean length of x, denoted $\|x\|$, is

$$\|x\| = \sqrt{x^* x} = \left(\sum_{i=1}^{m} \overline{x}_i x_i \right)^{\frac{1}{2}}$$

The cosine of the angle α between x and y is

$$\cos \alpha = \frac{x^* y}{\|x\| \, \|y\|}$$

The inner product is bilinear, that is,

$$\begin{aligned} (x_1 + x_2)^* y &= x_1^* y + x_2^* y \\ x^* (y_1 + y_2) &= x^* y_1 + x^* y_2 \\ (\alpha x)^* (\beta y) &= \overline{\alpha} \beta x^* y \end{aligned}$$

Orthogonal vectors

A pair of vectors x and y are orthogonal if $x^* y = 0$.

If x and y are real, this means that they lie at right angles to each other.

Two sets of vectors X and Y are orthogonal if every $x \in X$ is orthogonal to every $y \in Y$.

A set of nonzero vectors S is orthogonal if its elements are pairwise orthogonal. That is, if $x, y \in S$ and $x \neq y$, then $x^* y = 0$.

A set of nonzero vectors S is orthonormal if it is orthogonal and every $x \in S$ has $\|x\| = 1$.

Theorem: The vectors in an orthogonal set S are linearly independent.
Proof.

- if vectors in S are not independent, then some $v_k \in S$ can be expressed as a linear combination of other vectors

$$v_k = \sum_{i=1, i\neq k}^{n} c_i v_i$$

- recall that $\|v_k\|^2 = v_k^* v_k$, we have

$$\|v_k\|^2 = v_k^* v_k = v_k^* \Big(\sum_{i=1, i\neq k}^{n} c_i v_i \Big) = \sum_{i=1, i\neq k}^{n} c_i (v_k^* v_i) = 0$$

- this leads to a contradiction

Components of a vector

Note: inner products can be used to decompose arbitrary vectors into orthogonal components

Suppose $\{q_1, q_2, \cdots, q_n\}$ is an orthonormal set, and let v be an arbitrary vector. The vector

$$r = v - (q_1^* v)q_1 - (q_2^* v)q_2 - \cdots - (q_n^* v)q_n$$

is orthogonal to $\{q_1, q_2, \cdots, q_n\}$.

We see that v can be decomposed into $n+1$ orthogonal components

$$v = r + \sum_{i=1}^{n} (q_i^* v)q_i = r + \sum_{i=1}^{n} (q_i q_i^*)v$$

Remarks:

- If $\{q_i\}$ is a basis for $\mathbb{C}^m$, then $n = m$ and $r = 0$. We have

$$v = \sum_{i=1}^{m} (q_i^* v)q_i = \sum_{i=1}^{m} (q_i q_i^*)v$$

- two ways of writing the formula: first, linear combination of $\{q_i\}$; second, a sum of orthogonal projections $q_i q_i^*$ (very special rank one matrces)

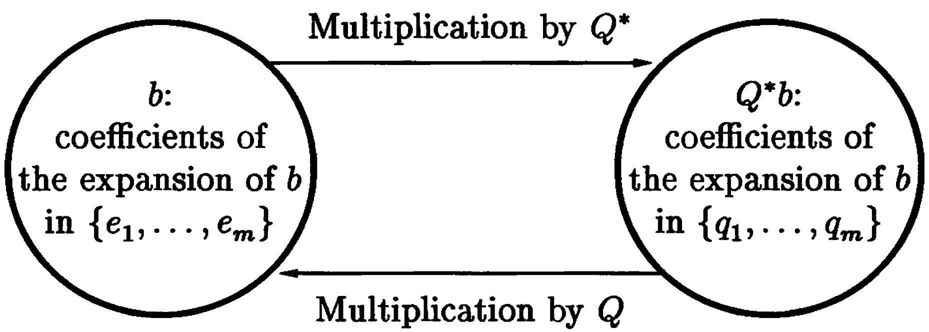
Unitary matrix

A square matrix $Q \in \mathbb{C}^{m\times m}$ is unitary if $Q^* = Q^{-1}$. That is, $Q^*Q = I$.

$$\begin{bmatrix} q_1^* \\ \hline q_2^* \\ \hline \vdots \\ \hline q_m^* \end{bmatrix} \begin{bmatrix} \left.\right| & \left.\right| & & \left.\right| \\ q_1 & q_2 & \cdots & q_m \\ \left.\right| & \left.\right| & & \left.\right| \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

Note that, $q_i^* q_j = \delta_{ij}$, where the symbol δ_{ij} is the Kronecker delta, equal to 1 if $i = j$ and 0 if $i \neq j$. Hence, the columns $\{q_i\}$ form an orthonormal basis for $\mathbb{C}^m$.

Note that, Qx is the linear combination of the columns of Q with coefficients x, and Q^*b is the vector of coefficients of the expansion of b in the basis of columns of Q. Schematically



Finally, multiplication by a unitary matrix Q preserves inner product and length:

$$(Qx)^*(Qy) = x^*Q^*Qy = x^*y \quad \text{and} \quad \|Qx\| = \|x\|$$

Lecture 3

A norm is a function $\|\cdot\| : \mathbb{C}^m \to \mathbb{R}$ that assigns a real-valued length to a vector.

A norm satisfies the following. For all vectors x, y and scalars $\alpha \in \mathbb{C}$:

$$\begin{aligned} &\|x\| \geq 0, \text{ and } \|x\| = 0 \text{ if and only if } x = 0 \\ &\|x + y\| \leq \|x\| + \|y\| \qquad (\textit{triangle inequality}) \\ &\|\alpha x\| = |\alpha| \, \|x\| \end{aligned}$$

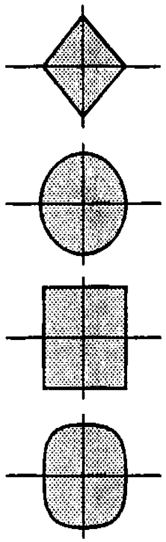
The following are common norms. The closed unit ball $\{x \in \mathbb{C}^m \ : \ \|x\| \leq 1\}$ are shown for $m = 2$ and real vectors.

$$\|x\|_1 = \sum_{i=1}^{m} |x_i|,$$

$$\|x\|_2 = \left(\sum_{i=1}^{m} |x_i|^2 \right)^{1/2} = \sqrt{x^* x},$$

$$\|x\|_\infty = \max_{1 \leq i \leq m} |x_i|,$$

$$\|x\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p} \quad (1 \leq p < \infty).$$



Another type of useful norms are the weighted p-norms.

Given any norm $\|\cdot\|$, a weighted norm can be defined as

$$\|x\|_W = \|Wx\|$$

where W can be any nonsingular matrix.

Matrix norm

Let $A \in \mathbb{C}^{m\times n}$. Let $\|\cdot\|_{(n)}$ and $\|\cdot\|_{(m)}$ be norms in $\mathbb{C}^n$ and $\mathbb{C}^m$. The induced matrix norm of A is the smallest number C such that

$$\|Ax\|_{(m)} \leq C\|x\|_{(n)}$$

holds for all $x \in \mathbb{C}^n$.

The matrix norm of A is denoted by $\|A\|_{(m,n)}$.

Equivalently, the norm of A can be computed by

$$\|A\|_{(m,n)} = \sup_{x\in\mathbb{C}^n,\ x\neq 0} \frac{\|Ax\|_{(m)}}{\|x\|_{(n)}} = \sup_{x\in\mathbb{C}^n,\ \|x\|_{(n)}=1} \|Ax\|_{(m)}$$

Example: the 1-norm of a matrix

Let $A \in \mathbb{C}^{m \times n}$, and consider the 1-norm for both $\mathbb{C}^n$ and $\mathbb{C}^m$. We compute the 1-norm $\|A\|_1$. Writing A in terms of columns

$$A = \left[\; \boldsymbol{a_1} \; \middle| \; \cdots \; \middle| \; \boldsymbol{a_n} \; \right]$$

Consider the set $\{x \in \mathbb{C}^n \; : \; \|x\|_1 = 1\}$. Any x in this set satisfies

$$\|Ax\|_1 = \|\sum_{j=1}^{n} x_j a_j\|_1 \le \sum_{j=1}^{n} |x_j| \, \|a_j\|_1 \le \left(\max_{1\le j\le n} \|a_j\|_1 \right) \sum_{j=1}^{n} |x_j| = \max_{1\le j\le n} \|a_j\|_1$$

This shows $\|A\|_1 \le \max_{1\le j\le n} \|a_j\|_1$.
Let k be the index which attains the $\max_{1\le j\le n} \|a_j\|_1$. Then

$$\|Ae_k\|_1 = \|a_k\|_1 = \max_{1\le j\le n} \|a_j\|_1$$

Thus,

$$\|A\|_1 = \sup_{x\in\mathbb{C}^n, \; \|x\|_1=1} \|Ax\|_1 \ge \|Ae_k\|_1 = \max_{1\le j\le n} \|a_j\|_1$$

Hence

$$\|A\|_1 = \max_{1\le j\le n} \|a_j\|_1$$

The 1-norm of A is the "*maximum column sum*" of A.

Cauchy-Schwarz and Holder inequalities

Let p, q satisfy $\frac{1}{p} + \frac{1}{q} = 1$, with $1 \leq p, q \leq \infty$.

The Holder inequality is

$$|x^* y| \leq \|x\|_p \, \|y\|_q$$

The Cauchy-Schwarz inequality is the special case when $p = q = 2$

$$|x^* y| \leq \|x\|_2 \, \|y\|_2$$

Example: 2-norm of an outer product

Let $A = uv^* \in \mathbb{C}^{m\times n}$ where $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$. We show $\|A\|_2 = \|u\|_2 \, \|v\|_2$.
For any $0 \neq x \in \mathbb{C}^n$

$$\|Ax\|_2 = \|uv^* x\|_2 = \|u\|_2 \, |v^* x| \leq \|u\|_2 \, \|v\|_2 \, \|x\|_2$$

So, $\|A\|_2 \leq \|u\|_2 \, \|v\|_2$.
On the other hand, take $x = v$, we have

$$\|Av\|_2 = \|uv^* v\|_2 = \|u\|_2 \|v\|_2^2$$

This shows

$$\|A\|_2 = \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \geq \frac{\|Av\|_2}{\|v\|_2} = \|u\|_2 \, \|v\|_2$$

General matrix norm

In general, a matrix norm must satisfy the following conditions:

$$\begin{aligned} &\|A\| \geq 0, \quad \text{and} \quad \|A\| = 0 \text{ if and only if } A = 0 \\ &\|A + B\| \leq \|A\| + \|B\| \\ &\|\alpha A\| = |\alpha| \, \|A\| \end{aligned}$$

One example is the Frobenius norm, which is not induced by a vector-norm. It is defined as

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}$$

Note that, we have

$$\|A\|_F = \sqrt{\mathrm{tr}(A^* A)} = \sqrt{\mathrm{tr}(A A^*)}$$

where $\mathrm{tr}(B)$ is the trace of B defined as the sum of its diagonal entries.

Invariance under unitary multiplication

Theorem: Let $A \in \mathbb{C}^{m\times n}$. Let $Q \in \mathbb{C}^{m\times m}$ be unitary. Then

$$\|QA\|_2 = \|A\|_2, \quad \text{and} \quad \|QA\|_F = \|A\|_F$$

Note that $\|Qy\|_2 = \|y\|_2$ for all y. Thus

$$\|QA\|_2 = \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|QAx\|_2}{\|x\|_2} = \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \|A\|_2$$

For the Frobenius norm, we have

$$\|QA\|_F = \sqrt{\operatorname{tr}(A^* Q^* QA)} = \sqrt{\operatorname{tr}(A^* A)} = \|A\|_F$$

Remarks:

- The above results hold if $Q \in \mathbb{C}^{p\times m}$, $p > m$, is a rectangular matrix with orthonormal columns
- Also hold if Q is multiplied on the right, and more generally, for rectangular Q with orthonormal rows

Lecture 4 - Singular value decomposition (SVD)

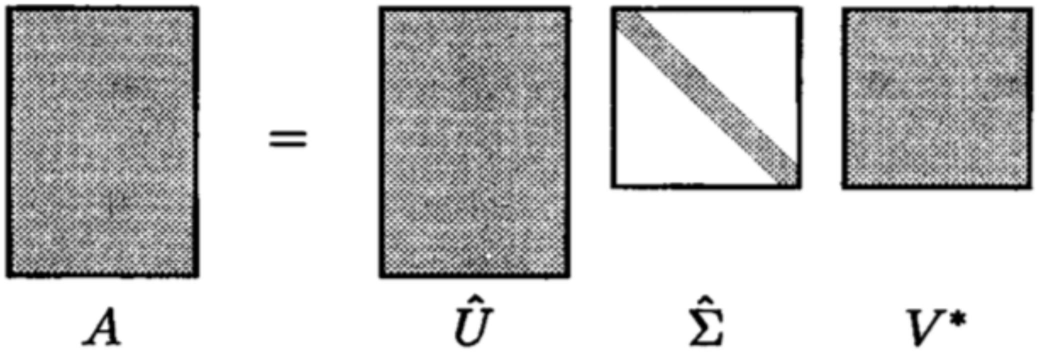
Let A be an $m \times n$ matrix $(m \geq n)$. We first present the reduced SVD:

$$A = \hat{U}\hat{\Sigma}V^*$$

where

- $\hat{\Sigma}$ is an $n \times n$ diagonal matrix
- $\hat{U}$ is an $m \times n$ matrix with orthonormal columns
- V is an $n \times n$ matrix with orthonormal columns

Schematically



Recall that

$$A = \hat{U}\hat{\Sigma}V^*$$

Multiplying both sides by V, we have

$$AV = \hat{U}\hat{\Sigma}$$

Schematically

$$\left[\begin{array}{c} \\ \\ A \\ \\ \\ \end{array}\right]\left[\begin{array}{c|c|c|c} v_1 & v_2 & \cdots & v_n \end{array}\right] = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \cdots & u_n \end{array}\right]\left[\begin{array}{cccc} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{array}\right]$$

Notice that, we have

$$Av_j = \sigma_j u_j, \qquad j = 1, 2, \cdots, n$$

Definitions:

- σ_j are called singular values satisfying $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$
- v_j are called right singular vectors, and u_j are called left singular vectors

Full SVD

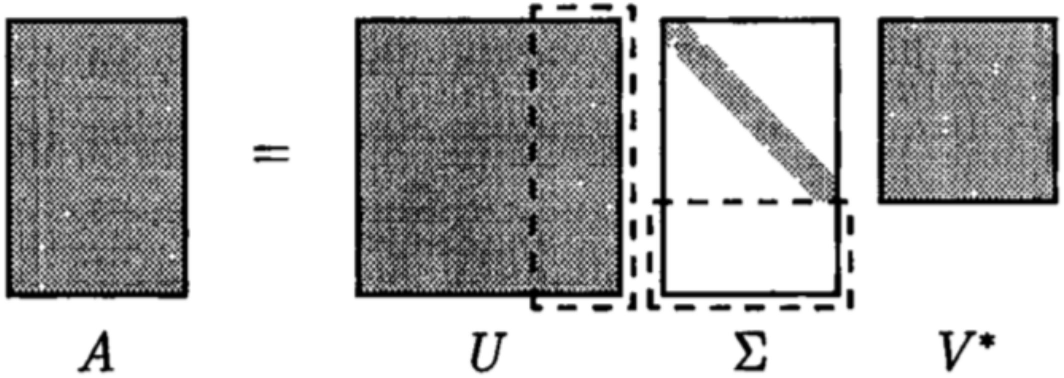
We state the full SVD (without the assumption of $m \geq n$):

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

where

- U is an $m \times m$ unitary matrix
- V is an $n \times n$ unitary matrix
- Σ is an $m \times n$ diagonal real matrix

Schematically (for the case $m \geq n$):



Note that, the singular values satisfy

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0, \qquad p = \min(m, n)$$

If A has rank $r \leq p$, then

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

Lecture 5

An observation: every matrix is diagonal when using a correct basis.

Let $A \in \mathbb{C}^{m \times n}$ and $A = U\Sigma V^*$. Let $x \in \mathbb{C}^n$. Consider the action of x under A: $b = Ax$.

We can represent the vector $x \in \mathbb{C}^n$ using the columns of V, namely

$$x = Vx'$$

Note that x' is the representation of x using the columns of V as basis.
Similarly, we can represent the vector $b \in \mathbb{C}^m$ using the columns of U, namely

$$b = Ub'$$

Note that b' is the representation of b using the columns of U as basis.

We have

$$b = Ax \iff b = U\Sigma V^* x \iff U^* b = \Sigma V^* x \iff b' = \Sigma x'$$

Using the new basis, the action of A becomes the action of a diagonal matrix.

Some properties

We summarize some properties related to SVD. Let $A \in \mathbb{C}^{m\times n}$ and $A = U\Sigma V^*$. Assume $\text{rank}(A) = r$.

- r is equal to the number of non-zero singular values.

  Pf: The rank of a diagonal matrix is equal to the number of nonzero entries on the diagonal.
- The space range(A) is spanned by the vectors $\{u_1, \cdots, u_r\}$.
- The space null(A) is spanned by the vectors $\{v_{r+1}, \cdots, v_n\}$.
- $\|A\|_2 = \sigma_1$, and $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$.

  Pf: First, we have $\|A\|_2 = \|\Sigma\|_2 = \max\{\sigma_j\} = \sigma_1$. Secondly, we have $\|A\|_F = \|\Sigma\|_F$, which gives the second property.
- The nonzero singular values of A are the square roots of the nonzero eigenvalues of A^*A or AA^*.

  Pf: Note that

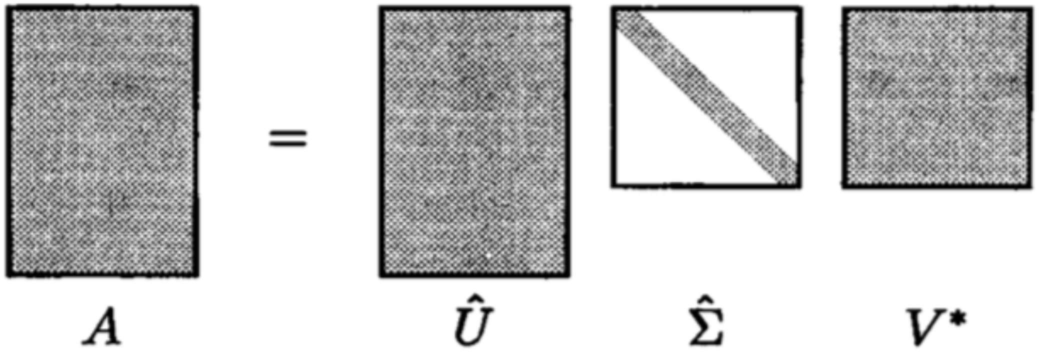
  $$A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^* U^* U\Sigma V^* = V(\Sigma^*\Sigma)V^*$$

  The two matrices A^*A and $\Sigma^*\Sigma$ have the same eigenvalues. The eigenvalues of $\Sigma^*\Sigma$ are $\sigma_1^2, \cdots, \sigma_p^2$ $(p = \min(m, n))$. If $n > p$, the remaining eigenvalues of $\Sigma^*\Sigma$ are zero.

Another representation of SVD

Let $A \in \mathbb{C}^{m \times n}$ and $A = U\Sigma V^*$ be its SVD. Assume $\text{rank}(A) = r \leq p = min(m, n)$.

Using $\text{rank}(A) = r$:



where $\hat{\Sigma}$ is $r \times r$, and $\hat{U}$ is $m \times r$ and V^* is $r \times n$.

The above is equivalent to

$$A = \sum_{j=1}^{r} \sigma_j \, u_j \, v_j^*$$

That is, A is a sum of rank one matrices.

Low-rank approximation

Recall

$$A = \sum_{j=1}^{r} \sigma_j \, u_j \, v_j^*$$

Let $0 \leq \nu < r$. We define

$$A_\nu = \sum_{j=1}^{\nu} \sigma_j \, u_j \, v_j^*$$

Note that $\text{rank}(A_\nu) = \nu$.

Theorem: Let $0 \leq \nu < r$. We have

$$\|A - A_\nu\|_2 = \inf_{B \in \mathbb{C}^{m \times n},\ \text{rank}(B) \leq \nu} \|A - B\|_2 = \sigma_{\nu+1}$$

Interpretation: the matrix A_ν gives the best rank-ν approximation of A.

Remark: similar result holds for the Frobenius norm.

Recall the theorem: Let $0 \le \nu < r$. We have

$$\|A - A_\nu\|_2 = \inf_{B \in \mathbb{C}^{m\times n},\ \mathrm{rank}(B) \le \nu} \|A - B\|_2 = \sigma_{\nu+1}$$

Proof.
Assume there is a matrix $B \in \mathbb{C}^{m\times n}$ such that

$$\mathrm{rank}(B) \le \nu, \quad \text{and} \quad \|A - B\| < \sigma_{\nu+1}$$

There is a space $W \subset \mathbb{C}^n$ of dimension $n - \nu$ such that

$$Bw = 0, \quad \text{for all } w \in W$$

Then, for any $w \in W$, we have

$$\|Aw\|_2 = \|(A - B)w\|_2 \le \|A - B\|_2 \|w\|_2 < \sigma_{\nu+1}\|w\|_2$$

Consider $S = \mathrm{span}\{v_1, \cdots, v_{\nu+1}\}$. Then $S \subset \mathbb{C}^n$ has dimension $\nu + 1$, and for any $s \in S$,

$$\|As\|_2 \ge \sigma_{\nu+1}\|s\|_2, \quad \text{since } As = \sum_{j=1}^{r} \sigma_j\, u_j\, v_j^* s = \sum_{j=1}^{\nu+1} \sigma_j\, u_j\, v_j^* s$$

Finally, we note that, since the sum of the dimensions of W and S is $n + 1$, there must be a nonzero vector lying in both W and S. This is a contradiction.