LA2 6

KYB

Thrn, it's a Fact

mathrnfact@gmail.com

October 16, 2020

Overview

8.1 Introduction to the SVD

Observation

- Recall that if $A \in \mathbb{C}^{n\times n}$ is Hermitian, A can be written as $A = VDV^*$ where $V \in \mathbb{C}^{n\times n}$ is unitary and $D \in \mathbb{R}^{n\times n}$ is diagonal.
- For any $A \in \mathbb{C}^{m\times n}$, $(AA^*)^* = AA^*$, that is, AA^* and A^*A are always Hermitian.

Then for $A \in \mathbb{C}^{n\times n}$ we can write A^*A by $A^*A = VDV^*$. Moreover for any vector $x \in \mathbb{C}^n$,

$$\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle \geq 0.$$

Thus every diagonal entry of D is nonnegative and we can find the square root matrix of D, say Σ,

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

We guess $A = V\Sigma V^*$. However, that is not true in general, because the matrix of the form is always Hermitian.

The SVD, Step 1

Suppose A is nonsingular. Then $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle = 0$ implies $x = 0$. Thus A^*A is positive definite.
Let $V = [v_1|\cdots|v_n]$, $D = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_n^2)$ where $\sigma_i > 0$. And define $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_n)$.

$$\langle Av_i, Av_j\rangle = \langle A^*Av_i, v_j\rangle = \left\langle \sigma_i^2 v_i, v_j\right\rangle = \sigma_i^2\langle v_i, v_j\rangle = \sigma_i^2\delta_{ij}$$

Define $u_i = \sigma_i^{-1}Av_i$, and then

$$AV = \left[Av_1 \mid \cdots \mid Av_n\right] = \left[\sigma_1 u_1 \mid \cdots \mid \sigma_n u_n\right] = U\Sigma$$

Note that

$$\langle u_i, u_j\rangle = \left\langle \sigma_i^{-1}Av_i, \sigma_j^{-1}Av_j\right\rangle = \frac{1}{\sigma_i\sigma_j}\langle A^*Av_i, v_j\rangle = \frac{\sigma_i^2}{\sigma_i\sigma_j}\delta_{ij} = \delta_{ij}.$$

Hence U is also unitary.

The SVD, Step 1

Rearrange σ_i's so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$. In this case, we say

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

is the Singular Value Decomposition of A.

The SVD, Step 2

Now we suppose A is singular and let the nullity be $n-r$, or

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Let $u_i = c$ for $i = 1, \cdots, r$. On the other hand for $i = r+1, \cdots, n$, $Av_i = 0$ because

$$\langle Av_i, Av_i \rangle = \langle A^* Av_i, v_i \rangle = \sigma_i^2 \langle v_i, v_i \rangle = 0,$$

Nevertherless, we can find $u_i's$ for $i = r+1, \cdots, n$ so that $\{u_i\}$ are orthonormal. Then

$$AV = \left[Av_1 \mid \cdots \mid Av_n \right] = \left[\sigma_1 u_1 \mid \cdots \mid \sigma_n u_n \right] = U\Sigma$$

Hence $A \in \mathbb{C}^{n \times n}$ always has the SVD.

8.2 The SVD for general matrices

The SVD, Step 3

The last step is that $A \in \mathbb{C}^{m\times n}$ has the SVD.
Assume $m \geq n$. Then we can find the diagonal matrix $D \in \mathbb{R}^{n\times n}$ and unitary matrix $V \in \mathbb{C}^{n\times n}$ such that $A^*A = VDV^*$. The diagonal entries of D are $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0$.
Define $u_i \in \mathbb{C}^m$ for $i = 1, \cdots, r$ so that $u_i = \sigma_i^{-1} A v_i$ and extend $\{u_1, \cdots, u_r\}$ to $\{u_1, \cdots, u_m\}$ which is orthonormal basis for $\mathbb{C}^m$.
Finally, define $\Sigma \in \mathbb{C}^{m\times n}$ by

$$\Sigma_{ij} = \sigma_i \delta_{ij}.$$

Then we get

$$AV = U\Sigma, \text{ or } A = U\Sigma V^*$$

If $n > m$, $A^* \in \mathbb{C}^{n\times m}$ and this is the above case.

The SVD

We can write $A \in \mathbb{C}^{m \times n}$ as simple as possible by using the SVD as follows:

Find $U \in \mathbb{C}^m$ and $V \in \mathbb{C}^n$ and $\Sigma \in \mathbb{R}^{m \times n}$ so that $A = U\Sigma V^*$. Let r be the largest index such that $\sigma_i > 0$ and define $\Sigma_1 \in \mathbb{R}^{r \times r}$ by $\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$ and split $U = [U_1|U_2]$ and $V = [V_1|V_2]$ by $U_1 = [u_1|\cdots|u_r]$, $U_2 = [u_{r+1}|\cdots|u_m]$, $V_1 = [v_1|\cdots|v_r]$ and $V_2 = [v_{r+1}|\cdots|v_n]$. Then

$$A = U\Sigma V^* = [U_1 \mid U_2] \left[\begin{array}{c|c} \Sigma_1 & 0 \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c} V_1^* \\ \hline V_2^* \end{array}\right] = U_1 \Sigma_1 V_1^*$$

We call $U_1 \Sigma_1 V_1^*$ the <u>reduced SVD</u> of A.

The outer product form of A

Let $A \in \mathbb{C}^{n\times n}$ with the reduced SVD $U_1\Sigma_1 V_1^*$ of rank r. For given $x \in \mathbb{C}^n$,

$$\begin{aligned} Ax = U_1\Sigma_1 V_1^* x &= U_1\Sigma_1 \begin{bmatrix} \langle x, v_1\rangle \\ \vdots \\ \langle x, v_r\rangle \end{bmatrix} \\ &= \begin{bmatrix} u_1 \mid \cdots \mid u_r \end{bmatrix} \begin{bmatrix} \sigma_1\langle x, v_1\rangle \\ \vdots \\ \sigma_r\langle x, v_r\rangle \end{bmatrix} \\ &= \sum_{i=1}^r \sigma_i \langle x, v_i\rangle u_i = \left(\sum_{i=1}^r \sigma_i u_i \otimes v_i\right) x \end{aligned}$$

Hence $A = \sum_{i=1}^r \sigma_i u_i \otimes v_i$.

Summary

- Every $A \in \mathbb{C}^{m\times n}$ has the SVD, $U\Sigma V^*$.
- If $\operatorname{rank}(A) = r$, there are only r positive singular values and A has the reduced SVD with $\Sigma_1 \in \mathbb{C}^{r\times r}$, $U_1\Sigma_1 V_1^*$.

How to find the SVD

1. Compute (or guess) eigen pairs of A^*A (or AA^*).
2. Orthogonalize(need not orthonormalize) $\{v_1, \cdots, v_n\}$ and compute $u_i = Av_i$ for $i = 1, \cdots, r = \operatorname{rank}(A)$ (need not $u_i = \frac{1}{\sigma_i}Av_i$)).
3. Extend $\{u_1, \cdots, u_r\}$ and orthonormalize $\{v_1, \cdots, v_n\}$ and $\{u_1, \cdots, u_m\}$.

Ex 8.1.3

$$A = \begin{bmatrix} 2 & 4 & 2 \\ 2 & 4 & 2 \\ 1 & 1 & -3 \end{bmatrix}.$$

Find the SVD of A in both matrix and outer product form.

Note

Since the first and second row are linearly dependent, A is singular. So 0 is an singular value of A.

Proof of Ex 8.1.3

$$A^T A = \begin{bmatrix} 2 & 2 & 1 \\ 4 & 4 & 1 \\ 2 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 2 & 4 & 2 \\ 1 & 1 & -3 \end{bmatrix} = \begin{bmatrix} 9 & 17 & 5 \\ 17 & 33 & 13 \\ 5 & 13 & 17 \end{bmatrix}$$

$$\begin{aligned} p_{A^T A}(r) &= \begin{vmatrix} r-9 & -17 & -5 \\ -17 & r-33 & -13 \\ -5 & -13 & r-17 \end{vmatrix} \\ &= r(r-11)(r-48). \end{aligned}$$

Proof of Ex 8.1.3

λ	σ	v_i	$\|v_i\|$
48	$4\sqrt{3}$	$(1,2,1)$	$\sqrt{6}$
11	$\sqrt{11}$	$(1,1,-3)$	$\sqrt{11}$
0	0	$(7,-4,1)$	$\sqrt{31}$

Take $u_i = \frac{1}{\sigma_i} A \frac{v_i}{\|v_i\|}$.

u_1	$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$
u_2	$(0,0,1)$

Take $u_3 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$

Then $u_3 \cdot u_i = 0$ for $i = 1,2$. Finally, Take $V = [\frac{v_1}{\|v_1\|} | \frac{v_2}{\|v_2\|} | \frac{v_3}{\|v_3\|}]$ and $U = [u_1|u_2|u_3]$ and $\Sigma = \mathrm{diag}(4\sqrt{3}, \sqrt{11}, 0)$. Then $A = U\Sigma V^T$.

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4\sqrt{3} & 0 & 0 \\ 0 & \sqrt{11} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} & \frac{-3}{\sqrt{11}} \\ \frac{7}{\sqrt{31}} & \frac{-4}{\sqrt{31}} & \frac{1}{\sqrt{31}} \end{bmatrix}$$

Proof of Ex 8.1.3

$$u_1 \otimes v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & 0 & 0 \end{bmatrix}$$

$$u_2 \otimes v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} & \frac{-3}{\sqrt{11}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} & \frac{-3}{\sqrt{11}} \end{bmatrix}$$

$$A = \sigma_1 u_1 \otimes v_1 + \sigma_2 u_2 \otimes v_2 = \begin{bmatrix} 2 & 4 & 2 \\ 2 & 4 & 2 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & -3 \end{bmatrix}$$

Ex 8.1.5

Let A be the 2×3 matrix defined as $A = uv^T$, where

$$u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Find the SVD of A.

Proof

Let $u_1 = (1/\sqrt{5}, 2/\sqrt{5})$,$v_1 = (1/\sqrt{2}, 0, 1/\sqrt{2})$ and $\sigma_1 = \sqrt{10}$. Then $A = \sigma_1 u_1 \otimes v_1$. Take $u_2 = (-2/\sqrt{5}, 1/\sqrt{5})$, $v_2 = (-1/\sqrt{2}, 0, 1/\sqrt{2})$, $v_3 = (0, 1, 0)$ and $\sigma_2 = 0$. Then

$$A = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$$

Ex 8.1.7

Suppose $A \in \mathbb{R}^{n \times n}$ has orthogonal columns. Find the SVD of A.

Ex 8.1.11

Suppose $A \in \mathbb{C}^{n\times n}$ is invertible and $A = U\Sigma V^*$ is the SVD of A. Find the SVD of each of the following matrices:

(a) A^*

(b) A^{-1}

(c) A^{-*}

Ex 8.1.13

Let $A \in \mathbb{C}^{n\times n}$ be normal, and let $A = XDX^*$ be the spectral decomposition of A. Explain how to find the SVD of A from X and D.

Ex 8.2.2

Let

$$A = \begin{bmatrix} 3 & 1 \\ 1 & -1 \\ 1 & -1 \\ -1 & -3 \end{bmatrix}.$$

Find the SVD of A and orthonormal bases for the four fundamental subspace of A.

Proof of Ex 8.2.2

$$A^T A = \begin{bmatrix} 12 & 4 \\ 4 & 12 \end{bmatrix}.$$

Then $\sigma_1 = 4$, $v_1 = (1, 1)$ and $\sigma_2 = 2\sqrt{2}$, $v_2 = (1, -1)$. Take $u_1 = Av_1 = (4, 0, 0, -4)$ and $u_2 = Av_2 = (2, 2, 2, 2)$. Revalue $u_1 = (1, 0, 0, -1)$ and $u_2 = (1, 1, 1, 1)$. Put $u_3 = (0, 1, -1, 0)$ and $u_4 = (0, 0, 1, 0) - \frac{1}{4}(1, 1, 1, 1) + \frac{1}{2}(0, 1, -1, 0) = (-1/4, 1/4, 1/4, -1/4)$. Finally normalize v_i and u_j. Then the SVD of A is

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2\sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Ex 8.2.7

Let $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ be given, and define $A = uv^T$. What are the singular values of A? Explain how to compute a singular value decomposition of A.

Ex 8.2.8

Let $u \in \mathbb{R}^n$ have Euclidean norm one, and define $A = I - 2uu^T$. Find the SVD of A.

Ex 8.2.9

Let $A \in \mathbb{R}^{m \times n}$ be nonsingular. Compute

$$\min\{\|Ax\|_2 \ : \ x \in \mathbb{R}^n, \|x\|_2 = 1\},$$

and find the vector $x \in \mathbb{R}^n$ that gives the minimum value.

Ex 8.2.10

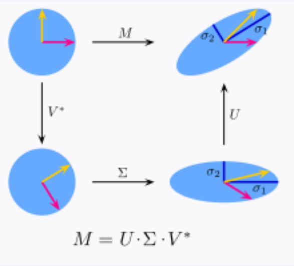
Let $A \in \mathbb{C}^{n \times n}$ be arbitrary. Using the SVD of A, show that there exist a unitary matrix Q and a Hermitian positive semidefinite matrix H such that $A = QH$. Alternatively, show that A can be written as $A = GQ$, where G is also Hermitian positive semidefinite and Q is the same unitary matrix. The decompositions $A = QH = GQ$ are the two forms of the polar decomposition of A.

Note

$$\det(A) = \det(Q)\det(H) = e^{i\theta} \cdot r$$

where $r = |\det(A)|$.

Geometrical meaning of the SVD



Link to Wiki : singular value decomposition

O'Neill - Elementary Differential Geometry

Ex 3.3.4 in O'Neill

Suppose $C \in \mathbb{R}^{3\times 3}$ is an orthogonal matrix. Then there is a number θ and an orthonormal sets $\{v_1, v_2, v_3\}$ such that

$$\begin{aligned} Cv_1 &= \cos\theta v_1 + \sin\theta v_2 \\ Cv_2 &= -\sin\theta v_1 + \cos\theta v_2 \\ Cv_3 &= \pm v_3 \end{aligned}$$

Note

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

rotates vectors in $\mathbb{R}^2$ counterclockwise throughout an angle θ with respect to the x axis.

Proof

$\det(rI - C)$ is a polynomial of degree 3. So it has at least one real solution, say λ_3. Let v_3 be the corresponding e.vec with norm one. Then

$$\lambda_3^2 = \lambda_3 v_3 \cdot \lambda_3 v_3 = Cv_3 \cdot Cv_3 = C^T Cv_3 \cdot c_3 = \|v_3\|^2 = 1.$$

So $\lambda_3 = \pm 1$. Now extend v_3 to an orthonormal basis $\{v_1, v_2, v_3\}$. Then

$$Cv_1 \cdot v_3 = \pm Cv_1 \cdot Cv_1 = \pm C^T Cv_1 \cdot v_3 = \pm v_1 \cdot v_3 = 0.$$

So $Cv_1 = a_1 v_1 + b_1 v_2$ and $Cv_2 = a_2 v_1 + b_2 v_2$. Since $\|Cv\| = \|v\|$, $a_1^2 + b_1^2 = 1$ and $a_2^2 + b_2^2 = 1$. $v_1 \cdot v_2 = 0$ implies $a_1 a_2 + b_1 b_2 = 0$. Then we can find (if you need, interchange v_1 and v_2 each other) some θ such that

$$\begin{aligned} a_1 &= \cos\theta, \quad a_2 = \sin\theta \\ b_1 &= -\sin\theta, \quad b_2 = \cos\theta. \end{aligned}$$

Ex 8.2.11

Let $A \in \mathbb{C}^{n\times n}$. Prove that $\|Ax\|_2 \le \sigma_1 \|x\|_2$ for all $x \in \mathbb{C}^n$, where σ_1 is the largest singular value of A.

Ex 8.2.12

Let $A \in \mathbb{C}^{n\times n}$. Prove that $\|Ax\|_2 \ge \sigma_n \|x\|_2$ for all $x \in \mathbb{C}^n$, where σ_n is the smallest singular value of A.

Ex 8.2.13

Given $A \in \mathbb{C}^{m\times n}$, the pseudoinverse $A^\dagger \in \mathbb{C}^{n\times m}$ is defined by the condition that $x = A^\dagger b$ is the minimum-norm least-squares solution to $Ax = b$.

(a) Let $\Sigma \in \mathbb{C}^{m\times n}$ be a diagonal matrix. Find $\Sigma^\dagger$.

(b) Find the pseudoinverse of $A \in \mathbb{C}^{m\times n}$ in terms of the SVD of A.

Ex 8.2.14

Let $m > n$ and suppose $A \in \mathbb{R}^{m\times n}$ has full rank. Let the SVD of A be $A = U\Sigma V^T$.

(a) Find the SVD of $A(A^TA)^{-1}A^T$.

(b) Prove that $\left\|A(A^TA)^{-1}A^Tb\right\|_2 \leq \|b\|_2$ for all $b \in \mathbb{R}^m$.

Ex 8.2.15

The *Frobenius norm* $\|\cdot\|_F$ on $\mathbb{C}^{m\times n}$ is defined by

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|A_{ij}|^2}.$$

Ex 8.2.15

(a) Prove that if $U \in \mathbb{C}^{m\times m}$ is unitary, then

$$\|UA\|_F = \|A\|_F.$$

Similarly, if $V \in \mathbb{C}^{n\times n}$ is unitary, then

$$\|AV\|_F = \|A\|_F.$$

Ex 8.2.15

(b) Let $A \in \mathbb{C}^{m\times n}$ be given, and let $r > 0$ such that $r < \operatorname{rank}(A)$. Find $B \in \mathbb{C}^{m\times n}$ of rank r such that B solve

$$\begin{aligned} &\min \|A - B\|_F \\ &\text{s.t.}\ \operatorname{rank}(B) = r. \end{aligned}$$

Solving least-squares problems using the SVD

Recall

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$

- $x \in \mathbb{R}^n$ be a LS solution to $Ax = b \Leftrightarrow x$ is a solution to $A^TAx = A^Tb$.
- $x \in \mathbb{R}^n$ be a MN-LS solution to $Ax = b \Leftrightarrow x$ is a LS sol to $Ax = b$ and $x \in \mathrm{col}(A^T)$.
- A has the SVD such that $A = U\Sigma V^T$.
- $A^\dagger = V\Sigma^\dagger U^T$.

If $U \in \mathbb{R}^{n\times n}$ is orthogonal and $x \in \mathbb{R}^n$,

- $\|Ux\|_2^2 = Ux \cdot Ux = U^TUx \cdot x = x \cdot x = \|x\|_2^2$.

MN-LS solution using the SVD, Step 1

$A = U\Sigma V^T$, x^* : LS solution to $Ax = b$.

$$\|Ax - b\|_2^2 = \left\|U\Sigma V^T x - b\right\|_2^2 = \left\|U\Sigma V^T x - UU^T b\right\|_2^2 = \left\|\Sigma V^T x - U^T b\right\|_2^2$$

Write $y = V^T x$.

$$\|Ax - b\|_2^2 = \left\|\Sigma y - U^T b\right\|_2^2$$

Let $U_1 \Sigma_1 V_1$ be the reduced SVD of A and write $y = [w^T | z^T]^T$ where $w \in \mathbb{R}^r$.

$$\left\|\Sigma y - U^T b\right\|_2^2 = \left\| \begin{bmatrix} \Sigma_1 w \\ \hline 0 \end{bmatrix} - \begin{bmatrix} U_1^T b \\ \hline U_2^T b \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \Sigma_1 w - U_1^T b \\ \hline -U_2^T b \end{bmatrix} \right\|_2^2$$

So we get

$$\|Ax - b\|_2^2 = \left\|\Sigma_1 w - U_1^T b\right\|_2^2 + \left\|U_2^T b\right\|_2^2.$$

MN-LS solution using the SVD, Step 2

By Step 1, if x^* is a LS solution to $Ax = b$, $x^* = Vy^*$ where $y^* = [w^{*T}|z^{*T}]^T$ and $w^* \in \mathbb{R}^r$ is a LS solution to $\Sigma_1 w = U_1^T b$.

Note that 1) such w^* is unique and 2) $z^* \in \mathbb{R}^{n-r}$ is arbitrary.

Take MN-SL $\bar{x}$ to $Ax = b$, i.e.,

- $\hat{x}$ is a LS solution to $Ax = b$
- $\bar{x}_2^2 = \min\{\|x^*\|_2^2 : x^*$ is a LS to $Ax = b\}$.

If we write $\bar{x} = V_1\bar{w} + V_2\bar{z}$, $\bar{w} = w^*$ and

$$\|\bar{x}\|_2^2 = \|w^*\|_2^2 + \|\bar{z}\|_2^2 \leq \|w^*\|_2^2 + \|z^*\|_2^2.$$

Since z^* is arbitrary, $\bar{z} = 0$, or $\bar{x} = V_1 w^*$.

MN-LS solution using the SVD, Step 3

Since Σ_1 is invertible, $w^* = \Sigma_1^{-1} U_1^T b$.

$$\begin{aligned} \bar{x} &= V_1 w^* = V_1 \Sigma_1^{-1} U_1^T b = (U_1 \Sigma_1 V_1^T)^\dagger b \\ &= (U \Sigma V^T)^\dagger b = A^\dagger b. \end{aligned}$$

Hence $A^\dagger = V \Sigma^\dagger U^T$.

Ex 8.3.4

Suppose $A \in \mathbb{R}^{m \times n}$ has SVD $A = U\Sigma V^T$, and we write $U = [U_1 | U_2]$, where U_1 form a basis for $\operatorname{col}(A)$ and the columns of U_2 form a basis for $\mathcal{N}(A^T)$. Show that, for $b \in \mathbb{R}^m$, $U_2 U_2^T b$ is the projection of b onto $\mathcal{N}(A^T)$ and $\left\|U_2 U_2^T b\right\| = \left\|U_2^T b\right\|$.

Ex 8.3.7

Let $A \in \mathbb{R}^{m\times n}$ have rank r. Write the formular for the MN-LS solution to $Ax = b$ in outer product form.

The Smith normal form of a matrix

Recall

- $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields.
- $\mathbb{Z}$ is not a field but it is closed under addition, multiplication and additive inverse operator.
- LA6 - Equivalance Relation and Partition of Set

Definition

A matrix $U \in \mathbb{Z}^{n\times n}$ is called unimodular if its determinant is 1 or -1.

Link to SNF - FTFGAG

Theorem (368, The Smith normal form)

Lset $A \in \mathbb{Z}^{m\times n}$ be given. There exist unimodular matrices $U \in \mathbb{Z}^{m\times m}$, $V \in \mathbb{Z}^{n\times n}$ and a diagonal matrix $S \in \mathbb{Z}^{m\times n}$ such that $A = USV$, the diagonal entries of S are $d_1, \cdots, d_r, 0, \cdots, 0$, each $d_i > 0$ and $d_i | d_{i+1}$ for $i = 1, \cdots, r-1$. S is called the Smith normal form of A.

$d_1, \cdots, d_r$ are called the elementary divisors (or the invariant factors).

Ex 4.4.7

Suppose $A \in \mathbb{R}^{n \times n}$ is invertible and has integer entries, and assume $\det(A) = \pm 1$. Prove that A^{-1} also has integer entries.

By Ex 4.4.7, U^{-1} and V^{-1} belong to $\mathbb{Z}^{m \times m}$ and $\mathbb{Z}^{n \times n}$, respectively. Define $W = V^{-1}$ and then

$$A = USV = USW^{-1}$$

The Division Algorithm

Let $m, n \in \mathbb{Z}$ with $m > n$. Then there is $q, r \in \mathbb{Z}$ such that $0 \leq r < n$ and

$$m = qn + r.$$

Elementary Matrices

For given $\lambda(\neq 0) \in \mathbb{R}$ and $i, j = 1, \cdots, n$, consider the following $n \times n$ matrices

- $M_{ij}(\lambda) : e_i \mapsto e_i + \lambda e_j$ for $k = i$; otherwise $e_k \mapsto e_k$.
- $A_{ij} : e_i \leftrightarrow e_j$ and $e_k \mapsto e_k$ for $k \neq i, j$.
- $N_i(\lambda) : e_i \mapsto \lambda e_i$ and $e_k \mapsto e_k$ for $k \neq i$.

Check

- $\det(M_{ij}(\lambda)) = 1$, $\det(A_{ij}) = -1$, and $\det(N_i(\lambda)) = \lambda$.
- $M_{ij}(\lambda)^{-1} = M_{ij}(-\lambda)$, $A_{ij}^{-1} = A_{ij}$, and $N_i(\lambda)^{-1} = N_i(\lambda^{-1})$.

Example

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 1 \\ 3 & 5 & 4 & 1 \\ 4 & 7 & 8 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 4 & 1 \\ 2 & 3 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 4 & 7 & 8 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 1 \\ 3 & 5 & 4 & 1 \\ 4 & 7 & 8 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 1 \\ 5 & 7 & 6 & 3 \\ 4 & 7 & 8 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 1 \\ 3 & 5 & 4 & 1 \\ 4 & 7 & 8 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 6 & 4 & 2 \\ 3 & 5 & 4 & 1 \\ 4 & 7 & 8 & 1 \end{bmatrix}$$

The Smith normal form, Step 1

Given $X \in \mathbb{Z}^{m\times n}$, multiply A_{ij}(or perform row and column interchanges) so that X_{11} is the smallest nonzero absolute value in all entries of X.

The Smith normal form, Step 2

If $X_{11}|X_{1j}$ for each $j = 2, \cdots, n$ go to Step 3. Otherwise, take the smallest value of j such that $X_{11}|X_{1j}$ fails. Then by Euclidean Algorithm, we can find q, r such that $X_{1j} = qX_{11} + r$. Mulitply replace X to $M_{j1}(-q)X$ and Go to Step 1.

The Smith normal form, Step 3

If $X_{11}|X_{i1}$ for each $i = 2, \cdots, m$, go to Step 4. Otherwise, take the smallest value of i such that $X_{11}|X_{i1}$ fails. Choose q, r, and replace X to $XM_{i1}(-q)$ and go to Step 1.

The Smith normal form, Step 4

Now X_{11} divides other entries in the first row and column. Add muliples of column 1 to culumns $2, 3, \cdots, n$, to zero out those entries. Similarly, add mutiples of row 1 to rows $2, 3, \cdots, m$ to zero out those entries. Then we get

$$X = \left[\begin{array}{c|c} X_{11} & 0 \\ \hline 0 & \tilde{X} \end{array}\right],$$

where $\tilde{X}$ is $(m-1) \times (n-1)$.

The Smith normal form, Step 5

If $m - 1 - 0$ or $n - 1 = 0$, then X is diagonal; otherwise aplly Steps 1 through 4 to the submatrix $\tilde{X}$.

The Smith normal form, Step 6

Now X is diagonal. Rearrange X_{ii}'s so that $0 < X_{11} \leq X_{22} \leq \cdots \leq X_{rr}$ and $X_{r+1,r+1} = \cdots = 0$. If there is $i \leq r-1$ such that $X_{ii} \not| X_{jj}$ for some $j > i$, add row j to row i and apply Step 1. If $X_{11}|X_{22}|\cdots|X_{rr}$, stop.
For each step, we just multiplied $P_s \in \mathbb{Z}^{m\times m}$ and $Q_s \in \mathbb{Z}^{n\times n}$ to X, P_sXQ_t. Thus

$$S = P_kP_{k-1}\cdots P_1XQ_1Q_2\cdots Q_l$$

Note that each P_t and Q_s are of the types $M_{ij}(\lambda)$ or A_{ij}. Hence $P_kP_{k-1}\cdots P_1$ and $Q_1Q_2\cdots Q_l$ are unimodular, as desired.

Ex 8.5.2

Let

$$A = \begin{bmatrix} 8 & 4 & 16 \\ 10 & 5 & 20 \\ 11 & 7 & 7 \end{bmatrix}.$$

Find the Smith decomposition $A = USV$ of A.

Application of the Smith normal form

Recall $(\mathbb{Z}_p, +, \cdot)$ is a ring, and it is a field if and only if p is a prime number.(LA1, LA6)

Theorem (372)

Let $A \in \mathbb{Z}^{n\times n}$, and let $\tilde{A} \in \mathbb{Z}_p^{n\times n}$ be obtained by replacing each entry of A by its congruence class modulo p. Then the congruence class of $\det(A)$ modulo p is the same as the $\det(\tilde{A})$ in $\mathbb{Z}_p$.

$$\det(A) \equiv \det(\tilde{A}) \textit{ mod } p$$

Corollary (373)

$\tilde{A}$ is singular if and only if $p | \det(A)$.

Definition

Let $A \in \mathbb{Z}^{n \times n}$. The p-rank of A is the rank of $\tilde{A}$.

Theorem (375)

1. *Let $A \in \mathbb{Z}^{n \times n}$ and let $S \in \mathbb{Z}^{n \times n}$ be the Smith normal form of A, with nonzero diagonal entries $d_1, \cdots, d_r$. Then the rank of A is r.*
2. *Let $B \in \mathbb{Z}_p^{n \times n}$ and let $T \in \mathbb{Z}_p^{n \times n}$ be the Smith normal form over $\mathbb{Z}_p$ of B with nonzero diagonal entries $e_1, \cdots, e_s$. Then the rank of B is s.*

Corollary (376)

Let $A \in \mathbb{Z}^{n \times n}$ and let $S \in \mathbb{Z}^{n \times n}$ be the Smith normal form of A, with nonzero diagonal entries $d_1, \cdots, d_r$. Let p be prime and let k be the largest integer such that p does not divide d_k. Then thee p-rank of A is k.

Remark 1

Suppose $A \in \mathbb{Z}^{n\times n}$ has the Smith normal form USV. Since $\det(U) = \dim(V) = 1$, $\dim(\tilde{U}) = \dim(\tilde{V}) = 1 \mod p$. So $\tilde{U}\tilde{S}\tilde{V}$ is the Smith normal form of $\tilde{A}$ over $\mathbb{Z}_p$.

Remark 2

If A has p-rank s, then so does A^T.

Example 377 in 8.5

$$A = \begin{bmatrix} 3 & 2 & 10 & 1 & 9 \\ 7 & 6 & 8 & 9 & 5 \\ -100 & -102 & -2 & -204 & 46 \\ -1868 & -1866 & 26 & -3858 & 1010 \\ -27204 & -27202 & 34 & -54734 & 13698 \end{bmatrix}$$

Find 5-rank of A.

Remark

Suppose $A \in \mathbb{Z}^{n\times n}$ has p-rank s. By theorem 375, the s

Proof

$$\tilde{A} = \begin{bmatrix} 3 & 2 & 0 & 1 & 4 \\ 2 & 1 & 3 & 4 & 0 \\ 0 & 3 & 3 & 1 & 1 \\ 2 & 4 & 1 & 2 & 0 \\ 1 & 3 & 4 & 1 & 3 \end{bmatrix} \in \mathbb{Z}_5^{5\times 5}.$$

Apply row operation.

$$\begin{bmatrix} 3 & 2 & 0 & 1 & 4 \\ 2 & 1 & 3 & 4 & 0 \\ 0 & 3 & 3 & 1 & 1 \\ 2 & 4 & 1 & 2 & 0 \\ 1 & 3 & 4 & 1 & 3 \end{bmatrix} \to \begin{bmatrix} 3 & 2 & 0 & 1 & 4 \\ 0 & 3 & 3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

So 5-rank of A is 3.

The End