7.4 Singular value decomposition

Let A be an $m \times n$ nonzero matrix.

Then $A^T A$ is $n \times n$ and symmetric.

$\{\vec{v}_1, \ldots, \vec{v}_n\}$: orthonormal eigvects for $A^T A$

$\lambda_1, \ldots, \lambda_n$: corresponding eigvals

Each $\lambda_i \geq 0$. Why?

$$0 \leq \|A\vec{v}_i\|^2 = (A\vec{v}_i)^T A\vec{v}_i = \vec{v}_i^{\,T} A^T A \vec{v}_i$$
$$= \vec{v}_i^{\,T}(\lambda_i \vec{v}_i) = \lambda_i \|\vec{v}_i\|^2 = \lambda_i$$

Renumber the eigvals of $A^T A$ so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

The <u>singular values of A</u> are $\sigma_i = \sqrt{\lambda_i}$

<u>Expls 1 & 2</u>

Find the singular values of $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$.

$$A^T A = \begin{pmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{pmatrix} \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$$

$$p(\lambda) = \cdots = -\lambda^3 + 450\lambda^2 - 32400\lambda$$
$$= -\lambda(\lambda - 360)(\lambda - 90)$$

eigvals of A^TA : $0, 360, 90$

order high to low:

$\lambda_1 = 360$, $\lambda_2 = 90$, $\lambda_3 = 0$

take square roots:

singular values of A

$$\boxed{\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0}$$

<u>Thm 9</u> Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ be the eigvals of A^TA. Suppose

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0, \quad \lambda_{r+1} = \cdots = \lambda_n = 0.$$

Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be orthonormal eigvects corresponding to $\lambda_1, \ldots, \lambda_n$.

not orthonormal

Then $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an orthogonal basis for $\operatorname{Col} A$, so $\operatorname{rank} A = r$.

$= \dim(\operatorname{Col} A)$

<u>Fact</u>: If $\vec{v}_i$ is a unit eigvect of A^TA corr. to λ_i, then $\sigma_i = \|A\vec{v}_i\|$. Why?

When we showed $\lambda_i \geq 0$, we calculated $\|A\vec{v}_i\|^2 = \lambda_i \|\vec{v}_i\|^2$. If $\|\vec{v}_i\| = 1$, then
$\|A\vec{v}_i\| = \sqrt{\lambda_i} = \sigma_i$.

Pf of Thm 9

$i \neq j \Rightarrow (A\vec{v}_i)^T(A\vec{v}_j) = \vec{v}_i^T A^T A \vec{v}_j = \vec{v}_i^T(\lambda_j \vec{v}_j)$

$= \lambda_j \underbrace{\cancel{(\vec{v}_i \cdot \vec{v}_j)}}_{0} = 0.$

$\therefore \{A\vec{v}_1, \dots, A\vec{v}_r\}$ is an orthogonal set. ✓

$1 \leq i \leq r \Rightarrow \|A\vec{v}_i\| = \sigma_i = \sqrt{\lambda_i} > 0$

$\Rightarrow A\vec{v}_i \neq \vec{0}$

$\therefore \{A\vec{v}_1, \dots, A\vec{v}_r\}$ is lin. indep. ✓

So $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ is an orthogonal basis for $H = \text{Span}\{A\vec{v}_1, \dots, A\vec{v}_r\}$. Need to show $H = \text{Col } A$.

Each $A\vec{v}_i \in \text{Col } A$, so $H \subset \text{Col } A$. ✓

Let $\vec{y} \in \text{Col } A$. Choose $\vec{x}$ such that $\vec{y} = A\vec{x}$.

Write $\vec{x}$ as $\vec{x} = c_1\vec{v}_1 + \dots + c_n\vec{v}_n$. Then

$$\vec{y} = A\vec{x} = c_1 A\vec{v}_1 + \dots + c_r A\vec{v}_r + c_{r+1} A\vec{v}_{r+1} + \dots + c_n A\vec{v}_n.$$

For all $i > r$, we have $\|A\vec{v}_i\| = \sigma_i = \sqrt{\lambda_i} = 0$, so $A\vec{v}_i = \vec{0}$. Thus,

$$\vec{y} = c_1 A\vec{v}_1 + \cdots + c_r A\vec{v}_r \in H.$$

Since $\vec{y}$ was arbitrary, this shows $\operatorname{Col} A \subset H$.
Combined with $H \subset \operatorname{Col} A$, we get $H = \operatorname{Col} A$. $\square$

Thm 10 (The singular value decomposition)

Let A be $m \times n$ with rank r (so that $r \le \min\{m,n\}$).
Let $\sigma_1, \ldots, \sigma_r$ be the first r singular values of A
(so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$).
Let Σ be the $m \times n$ matrix defined by

(A and Σ have the same dimensions)

call this $r \times r$ matrix D

$$\Sigma = (\sigma_1\vec{e}_1 \cdots \sigma_r\vec{e}_r\ \vec{0} \cdots \vec{0}) = \left(\begin{array}{c|c} \boxed{\begin{matrix} \sigma_1 & 0 \\ & \ddots \\ 0 & \sigma_r \end{matrix}} & 0 \\ \hline 0 & 0 \end{array}\right)$$

$m-r$ rows

$n-r$ cols

Then there exist orthogonal matrices (means orthonormal columns)
U $(m \times m)$ and V $(n \times n)$ such that

$$A = \underbrace{U \Sigma V^T}_{\text{the "singular value decomposition" (SVD) of } A}.$$

U and V are not unique

- The columns of U are called the <u>left singular vectors of A</u>.
- The columns of V are called the <u>right singular vectors of A</u>.

<u>Pf of Thm 10</u> Pay attention. Proof tells us how to solve SVD problems.

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$ be the eigvals of A^TA. Let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be orthonormal eigvects of A^TA corr. to $\lambda_1, \dots, \lambda_n$.

By Thm 9, $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ is an orthogonal basis for Col A.

in $\mathbb{R}^m$

$$\text{Normalize: } \vec{u}_i = \frac{1}{\|A\vec{v}_i\|} A\vec{v}_i = \frac{1}{\sigma_i} A\vec{v}_i \qquad (*)$$

Then $\{\vec{u}_1, \dots, \vec{u}_r\}$ is an orthonormal basis for Col A. Extend this to an orthonormal basis for $\mathbb{R}^m$, $\{\vec{u}_1, \dots, \vec{u}_m\}$. Define

$$U = (\vec{u}_1 \cdots \vec{u}_m), \quad V = (\vec{v}_1 \cdots \vec{v}_n),$$

so that both U and V are orthogonal matrices.

Check that this works:

$$AV = (\underbrace{A\vec{v}_1 \cdots A\vec{v}_r}_{\downarrow\ \text{by } (*)}\ \underbrace{A\vec{v}_{r+1} \cdots A\vec{v}_n}_{\downarrow\ \text{because for } i>r,\ \|A\vec{v}_i\| = \sigma_i = 0})$$

$$= (\overbrace{\sigma_1 \vec{u}_1 \cdots \sigma_r \vec{u}_r}\ \overbrace{\vec{0} \cdots \vec{0}}).$$

On the other hand,

$$U\Sigma = U(\sigma_1 \vec{e}_1 \cdots \sigma_r \vec{e}_r\ \vec{0} \cdots \vec{0})$$

$$= (\sigma_1 U\vec{e}_1 \cdots \sigma_r U\vec{e}_r\ \vec{0} \cdots \vec{0})$$

$$= (\sigma_1 \vec{u}_1 \cdots \sigma_r \vec{u}_r\ \vec{0} \cdots \vec{0}).$$

So $AV = U\Sigma$. But V is orthogonal, so $V^{-1} = V^T$. Therefore,

$$A = AVV^{-1} = AVV^T = U\Sigma V^T. \qquad \square$$

<u>Expl 3</u> Find a SVD of

$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}.$$

Follow the steps in theorem and proof.

Start w/ theorem

Find the eigenvalues and singular values.

Did that in Expls 1&2:

eigvals of A^TA in decr. order

$$\lambda_1 = 360, \quad \lambda_2 = 90, \quad \lambda_3 = 0$$

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

singular values of A

$r = 2$ (there are two positive singular values)

$$D = \begin{pmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{pmatrix}$$

Σ is 2×3 (same as A) w/ D in the top-left corner

$$\Sigma = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \checkmark$$

Need to find U and V. Look at proof.

Find orthonormal eigvects of A^TA, $\vec{v}_1, \vec{v}_2, \vec{v}_3$, corr. to $\lambda_1, \lambda_2, \lambda_3$.

$$\left.\begin{array}{l} (A^T A - \lambda_1 I \mid \vec{0}) \rightsquigarrow \vec{x} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \\ \\ (A^T A - \lambda_2 I \mid \vec{0}) \rightsquigarrow \vec{x} = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix} \\ \\ (A^T A - \lambda_3 I \mid \vec{0}) \rightsquigarrow \vec{x} = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \end{array}\right\}$$

automatically orthogonal since $A^T A$ is symmetric (Thm 3, the spectral thm)

Normalize: $\|\vec{x}_1\| = \sqrt{1+4+4} = \sqrt{9} = 3$

$\|\vec{x}_2\| = 3$, $\|\vec{x}_3\| = 3$

$$\vec{v}_1 = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ 1/3 \end{pmatrix}$$

$$V = \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix} \checkmark$$

Need to find U.

Aiming for $\underset{2\times 3}{A} = \underset{2\times 2}{U}\,\underset{2\times 3}{\Sigma}\,\underset{3\times 3}{V^T}$

$U = (\vec{u}_1, \vec{u}_2)$ need to find these

$r = 2$

$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1$, $\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2$

$$A\vec{v}_1 = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 4 + 22 + 28 \\ 8 + 14 - 4 \end{pmatrix}$$

$$= \frac{1}{3} \begin{pmatrix} 54 \\ 18 \end{pmatrix} = \begin{pmatrix} 18 \\ 6 \end{pmatrix} , \quad \sigma_1 = 6\sqrt{10}$$

$$\vec{u}_1 = \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix}$$

$$A\vec{v}_2 = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -8 - 11 + 28 \\ -16 - 7 - 4 \end{pmatrix}$$

$$= \frac{1}{3} \begin{pmatrix} 9 \\ -27 \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \end{pmatrix} , \quad \sigma_2 = 3\sqrt{10}$$

$$\vec{u}_2 = \begin{pmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{pmatrix}$$

Extend $\{\vec{u}_1, \vec{u}_2\}$ to an orthonormal basis of $\mathbb{R}^m = \mathbb{R}^2$.

Already a basis! Don't need to do anything.

$$U = \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{pmatrix} \checkmark$$

$$\boxed{A = \underbrace{\begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{pmatrix}}_{V^T}}$$

<u>Expl 4</u>

Find an SVD of $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & 2 \end{pmatrix}$.

$$\underset{3\times 2}{A} = \underset{3\times 3}{U} \, \underset{3\times 2}{\Sigma} \, \underset{2\times 2}{V^T}$$

$A^TA = \cdots = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$

eigvals of A^TA : . . . .

$\lambda_1 = 18$, $\lambda_2 = 0$ (decr. order)

singular values of A: $\sigma_1 = 3\sqrt{2}$, $\sigma_2 = 0$

r = # of positive sing. vals. $= 1$

$$\underset{(r\times r)}{D} = (3\sqrt{2})$$

Σ is 3×2 (same as A) with D in top left.

$$\Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

orthonormal eigvects of A^TA:

$\lambda_1 = 18 \quad \cdots \quad \vec{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ $\qquad V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$

$\lambda_2 = 0 \quad \cdots \quad \vec{v}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

$r = 1$

$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1$

$A\vec{v}_1 = \cdots = \begin{pmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{pmatrix}, \quad \sigma_1 = 3\sqrt{2}$

$\vec{u}_1 = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}$

Extend $\{\vec{u}_1\}$ to an orthonormal basis for $\mathbb{R}^3$.

How do I do this?

- Could pick $\vec{w}_2$ not a multiple of $\vec{u}_1$, then pick $\vec{w}_3$ not in Span$\{\vec{u}_1, \vec{w}_2\}$, then we use Gram-Schmidt, then normalize.

Alternative:

Want $\vec{w}_2, \vec{w}_3$ orthogonal to $\vec{u}_1$, i.e.

must solve $$\vec{u}_1^T \vec{x} = 0$$

$$\Updownarrow$$

$$\frac{1}{3}x_1 - \frac{2}{3}x_2 + \frac{2}{3}x_3 = 0$$

$$\Updownarrow$$

$$x_1 - 2x_2 + 2x_3 = 0$$

orthog. to $\vec{u}_1$ and lin. indep.

Solns:

$$\vec{x} = \begin{pmatrix} 2x_2 - 2x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \overset{\vec{w}_2}{\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}} + x_3 \overset{\vec{w}_3}{\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}}$$

$\{\underbrace{\vec{u}_1, \vec{w}_2}_{\text{already orthogonal}}, \vec{w}_3\}$ a basis for $\mathbb{R}^3$

Gram-Schmidt:

$$\vec{z}_3 = \vec{w}_3 - \overset{0\ (\text{orthog})}{\cancel{\frac{\vec{w}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}}}\, \vec{u}_1 - \frac{\vec{w}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2} \vec{w}_2$$

$$= \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \frac{-4}{4+1} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2/5 \\ 4/5 \\ 1 \end{pmatrix}$$

$\left\{ \vec{u}_1, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix} \right\}$: orthog. basis for $\mathbb{R}^3$

lengths: $\sqrt{5}$ $\quad \sqrt{4+16+24} = \sqrt{45}$

$$\vec{u}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{pmatrix}$$

$$\boxed{\begin{gathered} U = \begin{pmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{pmatrix}, \quad V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\ \Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A = U \Sigma V^T \end{gathered}}$$