Singular Value Decomposition

or, "what can you do if you can't diagonalize?"

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

Then A^t is an $n \times m$ matrix.

So A^tA is an $n \times n$ matrix,

and $(A^tA)^t = A^t(A^t)^t = A^tA$.

No matter what A is, A^tA is symmetric, and so is orthogonally diagonalizable!

But A^tA is not just symmetric!

Suppose λ is an eigenvalue for A^tA and $\vec{x}$ is an associated eigenvector.

$$\vec{x}^t(A^tA\vec{x}) = (\vec{x}^tA^t)\cdot(A\vec{x})$$
$$= (A\vec{x})^t(A\vec{x})$$
$$= \|A\vec{x}\|_2^2 \geq 0.$$

But also,

$$\vec{x}^t(A^tA\vec{x}) = \vec{x}^t(\lambda\vec{x})$$
$$= \lambda\vec{x}^t\vec{x}$$
$$= \lambda\|\vec{x}\|_2^2.$$

So

$$\lambda \|\vec{x}\|_2^2 = \vec{x}^t(A^t A\vec{x})$$

$$\lambda\|\vec{x}\|_2^2 = \|A\vec{x}\|_2^2$$

$$\lambda\|\vec{x}\|_2^2 \geq 0 \, .$$

Since $\|\vec{x}\|_2^2 > 0$, we can divide by $\|\vec{x}\|_2^2$ on both sides of the inequality

$\lambda \|\vec{x}\|_2^2 \geq 0$ to get

$$\boxed{\lambda \geq 0}$$

So all eigenvalues of $A^t A$ are non-negative real numbers!

Example 1:

Let $A = \begin{bmatrix} 1 & -3 & 6 \\ 0 & 5 & 4 \end{bmatrix}$.

$$A^t = \begin{bmatrix} 1 & 0 \\ -3 & 5 \\ 6 & 4 \end{bmatrix}$$

$$A^t A = \begin{bmatrix} 1 & 0 \\ -3 & 5 \\ 6 & 4 \end{bmatrix} \begin{bmatrix} 1 & -3 & 6 \\ 0 & 5 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -3 & 6 \\ -3 & 34 & 2 \\ 6 & 2 & 52 \end{bmatrix}$$

Using Wolfram Alpha, the eigenvalues are

$$\lambda = 0 \, , \; \frac{1}{2}\left(87 \pm \sqrt{349}\right)$$

$$0 \geq 0$$

$$\frac{1}{2}(87+\sqrt{349}) \geq 0$$

$\sqrt{349} < \sqrt{400} = 20$, so

$$\frac{1}{2}(87-\sqrt{349})$$

$$> \frac{1}{2}(87-20) > 0$$

So all eigenvalues of A^tA are non-negative!

Matrix Square Roots

Suppose D is a diagonal matrix,

$D = (D_{i,m})_{i,m=1}^{n}$. If

$D_{i,i} \geq 0$, then we can define

$\sqrt{D}$ to be the diagonal

matrix

$$(\sqrt{D})_{i,m} = \begin{cases} \sqrt{D_{i,i}}, & i = m \\ 0, & \text{otherwise.} \end{cases}$$

If A is diagonalizable,

$A = SDS^{-1}$, and $D_{ii} \geq 0$

for all $1 \leq i \leq n$, then we

can define

$$\boxed{\sqrt{A} = S\sqrt{D}S^{-1}}$$

Observe that

$$\sqrt{A} \cdot \sqrt{A} = (S\sqrt{D}S^{-1})(S\sqrt{D}S^{-1})$$

$$= S\sqrt{D}(\underbrace{\cancel{S^{-1}S}}_{I_n})\sqrt{D}S^{-1}$$

$$= S\sqrt{D} \cdot \sqrt{D}S^{-1}$$

$$= SDS^{-1}$$

$$= A$$

Back to Example 1:

$$A^tA = \begin{bmatrix} 1 & -3 & 6 \\ -3 & 34 & 2 \\ 6 & 2 & 52 \end{bmatrix}$$

is symmetric, and so orthogonally diagonalizable,

$$A^tA = ODO^t$$

where O is orthogonal and

$$D = \frac{1}{2}\begin{bmatrix} 87+\sqrt{349} & 0 & 0 \\ 0 & 87-\sqrt{349} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then

$$\sqrt{A^tA} = O\left(\frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{87+\sqrt{349}} & 0 & 0 \\ 0 & \sqrt{87-\sqrt{349}} & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)O^t$$

If A is $m \times n$, we define

$|A|$ ("absolute value of A") to

be

$$\boxed{|A| = \sqrt{A^tA}}$$

"The" SVD : special case

Suppose $|A|$ is invertible and set $\omega = A\,|A|^{-1}$.

Then

$$\omega \cdot |A| = A\,\underbrace{\cancel{|A|^{-1} \cdot |A|}}_{I_n}$$

$$\omega \cdot |A| = A.$$

Moreover,

$$\omega^t \omega = (A\,|A|^{-1})^t\,(A\,|A|^{-1})$$

$$\omega^t \omega = ((|A|^{-1})^t A^t)(A\,|A|^{-1})$$

$$w^t w = (|A|^t)^{-1}(A^t A)|A|^{-1}$$

since $A^t A$ is symmetric, so is $|A|$

$$w^t w = (|A|^{-1})(A^t A)|A|^{-1}$$

$$w^t w = |A|^{-1}(\sqrt{A^t A}\,\sqrt{A^t A})|A|^{-1}$$

$$w^t w = \cancel{|A|^{-1}}(\cancel{|A|} \cdot \cancel{|A|})\cancel{|A|^{-1}}$$

$$w^t w = I_n.$$

So w is an orthogonal matrix!

(not quite - w is not necessarily square)

Then since $A^t A$ is symmetric,

$A^t A$ is orthogonally diagonalizable

as $A^t A = O D O^t$.

Then

$$|A| = \sqrt{A^t A} = O\sqrt{D}\,O^t, \text{ so}$$

$$A = \omega|A| = \omega O\sqrt{D}\,O^t.$$

Now

$$\begin{aligned}(\omega O)^t \omega O &= O^t \omega^t \omega O \\ &= O^t O \\ &= I_n\end{aligned}$$

Rename

$$\begin{aligned}U &= \omega O \\ \Sigma &= \sqrt{D} \\ V &= O^t.\end{aligned}$$

Then

$$A = \omega |A|$$

$$A = \omega O \sqrt{D} O^t$$

$$\boxed{A = U \Sigma V}$$

Singular value decomposition of A

Here, Σ is diagonal,

V is orthogonal, and

$$U^t U = I_n.$$

Note U is $m \times n$, so can't be orthogonal if A is not square.

In General

If $|A|$ is not invertible,

then play the same game as

if it were for the $k \times k$

matrix $(k < n)$ consisting of

the nonzero eigenvalues of $A^t A$.

Then "pad" $\sqrt{D}$ with zeros,

and orthonormal columns or rows to

U and V, if necessary.

We again get

$$\boxed{A = U \Sigma V}$$

Try Wolfram Alpha!

$$A = \begin{bmatrix} 1 & -3 & 6 \\ 0 & 5 & 4 \end{bmatrix}$$

The SVD of A is horribly messy, but for Wolfram Alpha, U and V are orthogonal, but Σ is not a square matrix!

What we computed is the reduced SVD. Wolfram Alpha is computing the full SVD.

Full SVD

A is $m \times n$,

$$\boxed{A = \hat{U} \hat{\Sigma} \hat{V}}$$

where $\hat{U}$ is an $m \times m$ orthogonal matrix, $\hat{V}$ is an $n \times n$ orthogonal matrix, and $\hat{\Sigma}$ is an $m \times n$ matrix with $(\hat{\Sigma})_{i,k} = 0$ if $i \neq k$

for all $1 \leq i \leq m$, $1 \leq k \leq n$.

The "singular values" of A

are the diagonal entries of Σ (or $\hat{\Sigma}$ if you only want nonzero entries).