The Singular Value Decomposition

Linear Algebra
MATH 2076

The Matrix $M = A^TA$

What can we say about $M = A^TA$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^TA)^T = A^TA = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

Item (1) says that M has (real) eigenvalues $\lambda_1, \lambda_2 \ldots, \lambda_n$. Item (2) says that M has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$.

Let's examine

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

Here we find that

$$M = A^TA = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

We show that

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & \sqrt{3} & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

this is a so-called *singular value decomposition* for the matrix A.
We already saw that

$$A^T A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

So $\lambda_1 = 3, \lambda_2 = 2$ are eigenvalues for $A^T A$ with assoc'd unit eigenvectors $\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Note that $A\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $A\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ satisfy $\|A\vec{v}_1\| = \sqrt{3} = \sqrt{\lambda_1} = \sigma_1, \|A\vec{v}_2\| = \sqrt{2} = \sqrt{\lambda_2} = \sigma_2$, and $A\vec{v}_1 \perp A\vec{v}_2$.
Put

$$\vec{u}_1 = \frac{A\vec{v}_1}{\|A\vec{v}_1\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{u}_2 = \frac{A\vec{v}_2}{\|A\vec{v}_2\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Now choose $\vec{u}_3$ so that $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis of $\mathbb{R}^3$.

We have

$$\vec{u}_1 = \frac{A\vec{v}_1}{\|A\vec{v}_1\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{u}_2 = \frac{A\vec{v}_2}{\|A\vec{v}_2\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

We want to choose $\vec{u}_3$ so that $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis of $\mathbb{R}^3$.

One choice (there are only two possibilities) is $\vec{u}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

Finally, we see that $A = U\Sigma V^T$ where $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$ and $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix}$ are orthogonal matrices, and

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

That is,

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & \sqrt{3} & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The Matrix A^TA

We return to our discussion concerning A^TA where A is any $m \times n$ matrix? Evidently, A^TA is an $n \times n$ matrix, and $(A^TA)^T = A^TA$, so A^TA is *symmetric*. Thus:

1. A^TA has n real eigenvalues, say, $\lambda_1, \lambda_2 \dots, \lambda_n$.
2. A^TA is *orthogonally diagonalizable*.

Item (2) says that A^TA has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \dots, \vec{v}_n\}$ where, say, $A^TA\vec{v}_i = \lambda_i\vec{v}_i$.
Now look at

$$\begin{aligned} \|A\vec{v}_i\|^2 &= (A\vec{v}_i) \cdot (A\vec{v}_i) = (A\vec{v}_i)^T(A\vec{v}_i) = \vec{v}_i^T(A^TA)\vec{v}_i = \vec{v}_i^T\lambda_i\vec{v}_i \\ &= \lambda_i\vec{v}_i \cdot \vec{v}_i = \lambda_i\|\vec{v}_i\|^2 = \lambda_i\,. \end{aligned}$$

Thus $\lambda_i = \|A\vec{v}_i\|^2 \geq 0$. By relabeling, if necessary, we can assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$.

The Matrix $A^T A$

When A is an $m \times n$ matrix, $A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and there is an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $A^T A\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

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

Here $1 \leq r \leq n$ and when $r < n$, $A\vec{v}_{r+1} = \cdots = A\vec{v}_n = \vec{0}$. So, if $\vec{x} = \sum_{i=1}^{n} c_i \vec{v}_i$, then

$$A\vec{x} = \sum_{i=1}^{n} c_i A\vec{v}_i = \sum_{i=1}^{r} c_i A\vec{v}_i.$$

Thus, $\mathcal{CS}(A) = \mathcal{S}pan\{A\vec{v}_1, \ldots, A\vec{v}_r\}$. Is $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ LI?

A^TA where A is $m \times n$

- A^TA is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for A^TA
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with A^TA
- $A^TA\vec{v}_i = \lambda_i\vec{v}_i$ and $\lambda_i = \lambda_i\|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) are $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

Look at

$$A\vec{v}_i \cdot A\vec{v}_j = (A\vec{v}_i)^TA\vec{v}_j = \vec{v}_i^T(A^TA)\vec{v}_j = \lambda_j\vec{v}_i^T\vec{v}_j = \lambda_j\vec{v}_i \cdot \vec{v}_j = \lambda_j\delta_{ij}.$$

Thus $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an orthogonal set of non-zero vectors, so it is LI.
$\therefore$ $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$.

$A^T A$ where A is $m \times n$

- $A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for $A^T A$
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with $A^T A$
- $A^T A\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) are $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

It follows that $\text{rank}(A) = \dim \mathcal{CS}(A) = r$. Recall that $\mathcal{CS}(A)$ is vector subspace of $\mathbb{R}^m$.

For $1 \leq i \leq r$, let $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$. Then $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$. Add unit orthogonal vectors to get orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_r, \vec{u}_{r+1}, \ldots, \vec{u}_m\}$ for $\mathbb{R}^m$.

Matrices: A $(m \times n)$, $M = A^T A$ $(n \times n)$, and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) are $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0)$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Let $U = \left[\vec{u}_1 \ \vec{u}_2 \ldots \vec{u}_m\right]$ (U is $m \times m$) and $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$ (V is $n \times n$). Now define an $r \times r$ diagonal matrix D and an $m \times n$ matrix Σ by

$$\Sigma = \begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \text{where} \quad D = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ 0 & \vdots & \vdots & 0 \\ 0 & 0 & \ldots & \sigma_r \end{bmatrix}.$$

Then, $\boxed{A = U\Sigma V^T}$. This is a *singular value decomposition* for A.

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

1. Find an orthogonal diagonalization of $M = A^T A$.
2. Write down V and Σ.
3. Find $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$ and then U.

Find eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ of $M = A^T A$ and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (an orthon eigenbasis assoc'd with M, so a basis for $\mathbb{R}^n$).

Get $n \times n$ matrix $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$. Also, $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$, which gives D and then Σ.

Need $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$, an orthon basis for $\mathbb{R}^m$. Then have $m \times m$ matrix $U = \left[\vec{u}_1 \ \vec{u}_2 \ldots \vec{u}_m\right]$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}, A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ so $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

Need $\vec{u}_3 \cdot \vec{u}_1 = 0 = \vec{u}_3 \cdot \vec{u}_2$. Use $\vec{u}_3 = \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$. Then

$U = \frac{1}{3\sqrt{2}} \begin{bmatrix} 3 & 1 & -2\sqrt{2} \\ 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \end{bmatrix}$. $\frac{1}{6} \begin{bmatrix} 3 & 1 & -2\sqrt{2} \\ 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$