Section 7.4

Have seen a couple ways to diagonalize. Unfortunately, cannot diagonalize all matrices. But there is process that does something similar, for any matrix. The Singular Value Decomposition.

Skipped some sections in book so will skip the examples that lack context. Mostly concerned with mechanics.

Prelims

A ($m \times n$) may not be diagonalizable but note A^TA ($n \times n$) is symmetric, so this matrix is orthogonally diagonalizable.

So A^TA has n eigenvalues (may have repeats)

- Eigenvalues of A^TA are non-negative.

  Let $\vec{v}_i$ be eigenvector of A^TA. Note that $A\vec{v}_i$ is defined. Consider length squared of $(A\vec{v}_i)$

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

$= \lambda_i (\vec{v}_i \cdot \vec{v}_i)$

Since $\|A\vec{v}_i\|^2 \geq 0$ and $(\vec{v}_i \cdot \vec{v}_i) \geq 0$, must be that $\lambda_i \geq 0$ as well.

- Can arrange eigenvalues of A^TA in decreasing order so that

$$\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0$$

- Call the square roots of λ's the singular values of A (not A^TA) usually denoted as sigma, $\sigma_i = \sqrt{\lambda_i}$

- Since A^TA symmetric can assume/choose eigenvectors $\vec{v}_1, \dots \vec{v}_n$ to be orthonormal

Theorem: Suppose $\{\vec{v}_1, \dots \vec{v}_n\}$ is orthonormal basis of $\mathbb{R}^n$ made up of eigenvectors of A^TA, with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \lambda_r \dots \lambda_n \geq 0$. Let $\sigma_1, \dots \sigma_r$ be nonzero singular values of A. Then $\{A\vec{v}_1, \dots A\vec{v}_r\}$ is orthogonal basis for Col A and Rank $A = r$.

D

First, show Av_i are orthogonal:

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

$$= \vec{v}_i^T (A^T A \vec{v}_j)$$
$$= \vec{v}_i^T (\lambda_j \vec{v}_j)$$
$$= \lambda_j (\vec{v}_i \cdot \vec{v}_j)$$
$$= 0.$$

So $A\vec{v}_i$ are orthogonal!

(by the above)

Note $\|A\vec{v}_i\|^2 = (A\vec{v}_i) \cdot (A\vec{v}_i) = \lambda_i (v_i \cdot v_i)$
$$= \lambda_i \|\vec{v}_i\|^2$$

Thus we see
$$\|A\vec{v}_i\|^2 = \lambda_i \|\vec{v}_i\|^2$$
$$\|A\vec{v}_i\| = \sqrt{\lambda_i} \|\vec{v}_i\|$$
$$\|A\vec{v}_i\| = \sigma_i \|\vec{v}_i\|$$

Since $\vec{v}_i$ is nonzero vector, $\|A\vec{v}_i\|$ is 0 (meaning $A\vec{v}_i = \vec{0}$) if and only if singular value $\sigma_i = 0$ (so $i > r$, not one of nonzero singular values)

Since $A\vec{v}_i$ are orthogonal and nonzero vectors (for $i \le r$) they are linearly independent. And clearly each $A\vec{v}_i$ is in column space of A. $\{A\vec{v}_1, \dots A\vec{v}_r\}$ is linearly independent / orthogonal set in Col A. Does it span Col A?

Let $y \in \text{Col } A$. Then $y = A\vec{x}$ for some $\vec{x}$.
$\{\vec{v}_1, \dots \vec{v}_n\}$ is a basis for $\mathbb{R}^n$ so

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

So $\vec{y} = A(c_1\vec{v}_1 + \dots c_n\vec{v}_n)$

$= c_1 A\vec{v}_1 + \dots + c_n A\vec{v}_n$

But by above $Av_i = \vec{0}$ if $i > r$ so

$c_1 A\vec{v}_1 + \dots c_n A\vec{v}_n = c_1 A\vec{v}_1 + \dots c_r A\vec{v}_r$

Thus $\vec{y} = c_1 A\vec{v}_1 + \dots c_r A\vec{v}_r$

$\vec{y}$ is linear combination of $A\vec{v}_1 \dots A\vec{v}_r$.

They span Col A.

Thus $A\vec{v}_1 \dots A\vec{v}_r$ are orthogonal basis for Col A. Since there are r basis vectors, dimension of Col $A = r$, i.e. Rank $A = r$.

∎

SVD

$$A = \underbrace{U}_{m\times m}\underbrace{\Sigma}_{m\times n}\underbrace{V^T}_{n\times n}$$

Goal is to write A as

$$\underline{A} = U\Sigma V^T \quad ☆$$

where Σ is at least "close" to diagonal matrix.

Best we will be able to do is

$$\underbrace{\Sigma}_{\substack{m\times n \text{ matrix} \\ \text{same size as } A}} = \begin{array}{c} \\ \left[\begin{array}{cc} D & 0 \\ 0 & 0 \end{array}\right] \\ \begin{array}{cc} r \text{ columns} & n-r \text{ columns} \end{array} \end{array} \begin{array}{l} \} \, r \text{ rows} \\ \} \, m-r \text{ rows} \\ \\ \end{array}$$

where D is diagonal $r\times r$ matrix with nonzero entries

Theorem: Let A be $m \times n$ matrix with rank r. Then there exists $m \times n$ matrix Σ (of form above) where entries in D block are first r singular values of A (nonzero). There also exists $m \times m$ orthogonal U and $n \times n$ orthogonal V s.t.

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

Definition: Any decomposition of above form is a Singular Value Decomposition of A.

How to find SVD

<u>For $m \times n$ matrix A</u>: $\quad A = U \Sigma V^T$

Most of what we need is based off of $A^T A$ and its orthogonal diagonalization.

Step 1: Find orthogonal diagonalization of $A^T A$

$A^T A = P D P^T$ where D has decreasing values along diagonal

Step 2: Find singular values of A

$\sigma_1 = \sqrt{\lambda_1} \quad \sigma_2 = \sqrt{\lambda_2}$ etc.

Step 3: Make $m \times n$ matrix Σ. Top left corner is $r \times r$ diagonal matrix of nonzero σ_i's.

Step 4: Make V^T. This is same matrix on far right when we diagonalize A^TA.

$$A^TA = VDV^T$$

Step 5: Find U. For first r eigenvectors in V calculate $A\vec{v}_i$, make them unit vectors.

[If $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is orthogonal basis for $\mathbb{R}^m$, done. If not, extend $\{A\vec{v}_1 \ldots A\vec{v}_r\}$ until you have basis for $\mathbb{R}^m$.]

$$\begin{bmatrix} | & & | & | & | & | \\ A\vec{v}_1 & \cdots & A\vec{v}_r & & & \\ | & & | & | & | & | \end{bmatrix}$$

$m \times m$

$$\downarrow$$

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{array}{cc} Av_1 & Av_2 \end{array}$$