Week 15 Tuesday.

If A can be diagonalized. then we can find . P invertible s.t. $P^{-1}AP = D \iff A = P\cdot D\cdot P^{-1}$

If A is symmetric i.e. $A = A^T$. then we can take P to be orthogonal. then

$$A = \underset{\substack{\uparrow \\ \text{orthegonal}}}{P} \cdot \underset{\substack{\uparrow \\ \text{digonal}}}{D} \cdot \underset{\substack{\nwarrow \\ \text{orthogonal.}}}{P^T}$$

and $rk(A) = rk(D) =$ # of non-zero entries in the diagonal. because $rk(A) = rk(A \cdot \text{any invertible matrix})$

so $rk(A) = rk(P^{-1}A) = rk(P^{-1}\cdot A\cdot(P^T)^{-1}) = rk(D)$

= # of non-zero eigenvalues (with multiplicity).

None of these make sense for non-square matrix!

Say $A \in M_{m\times n}(\mathbb{R})$. $B = A^T\cdot A \in M_{n\times n}(\mathbb{R})$

Property of B:

1). B is symmetric $(A^T\cdot A)^T = A^T(A^T)^T = A^T A$

2) All eigenvalues of B is non-negative.

$$x^T\cdot Bx = x^T\cdot A^T\cdot A\cdot x = (Ax)^T\cdot(Ax) = \|Ax\|^2 \geq 0$$

So if $\vec{x}$ is an eigenvector w.r.t. λ then.

$$x^T\cdot Bx = x^T(\lambda\cdot x) = \lambda\cdot x^T\cdot x \Rightarrow \|Ax\|^2 = \lambda\cdot\|x\|^2. \Rightarrow. \lambda \geq 0$$

Therefore we can pick a set of eigenvectors of B. $\{\vec{v}_1,\cdots,\vec{v}_n\}$ that is orthonormal. order. $\vec{v}_1,\cdots,\vec{v}_n$ s.t.

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

Thm. $\{A\vec{v}_1, \cdots, Av_n\}$ pairwise orthogonal, and.

if $\lambda_1, \cdots, \lambda_r > 0$. then. $\{A\vec{v}_1, \cdots, A\vec{v}_r\}$ is linearly independet.

if $\lambda_r > 0$ & $\lambda_{r+1} = 0$ then. $\{A\vec{v}_1, \cdots, A\vec{v}_r\}$ forms an orthogonal basis for $Col(A)$.

Pf. If $i \neq j$ $(A\vec{v}_i)^T \cdot (A\vec{v}_j) = \vec{v}_i^T \cdot A^T \cdot A \cdot \vec{v}_j = \vec{v}_i^T \cdot \lambda_j \vec{v}_j = \lambda_j \cdot 0 = 0.$

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

if $\lambda_i > 0$ then $A\vec{v}_i \neq \vec{0}$.

[Given a set of orthogonal nonzero vectors, it must be linearly independent: $S = \{\vec{v}_1, \cdots, \vec{v}_r\}$ and.

$$\vec{u} = \Sigma \alpha_i \vec{v}_i = \vec{0}. \quad \text{then.} \quad \alpha_i = \vec{u} \cdot \frac{\vec{v}_i}{\|\vec{v}_i\|} = \vec{0}. \quad]$$

If there're exactly r eigenvalues $\lambda_i > 0$ then. $A\vec{v}_i = \vec{0}$ for $i > r$. so $span\{A\vec{v}_1, \cdots, A\vec{v}_n\} = span\{A\vec{v}_1, \cdots, A\vec{v}_r\}$

and. since $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$. we have

$$\underset{\overset{\|}{Col(A \cdot \underset{\overset{\|}{\left(\underset{v_1}{\downarrow} \cdots \underset{v_n}{\downarrow}\right)}}{V})}}{span\{A\vec{v}_1, \cdots, A\vec{v}_n\}} = Col(A).$$

So $\{A\vec{v}_1, \cdots, A\vec{v}_r\}$. span the $Col(A)$ and. are linearly independent. and are othogonal. so it gives an orthogonal basis for . $Col(A)$.

Singular Value Decomposition

Thm For $A \in M_{m\times n}(\mathbb{R})$ we have $A = \underset{m\times m}{U} \cdot \underset{m\times n}{\overset{\text{"close" to diagonal}}{\Sigma}} \cdot \underset{n\times n}{V^T}$ where

U, V^T: orthogonal

"close" to diagonal. $\Sigma_{ij}=0$ for $i \neq j$.

$$\begin{pmatrix} * & & & & \\ & * & & & \\ & & * & & \\ & & & * & \\ & & & & * \end{pmatrix}$$

Pf. $V^T = V^{-1}$ by orthogonality therefore it suffices to prove that $A \cdot V = U \cdot \Sigma$.

Taking $V = \begin{pmatrix} \downarrow & \downarrow & \cdots & \downarrow \\ v_1 & v_2 & & v_n \end{pmatrix}$ where $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is the set of orthonormal basis of $A^T \cdot A$ in the order of $\lambda_1 \geq \cdots \geq \lambda_n$.

$$A \cdot V = \Big(\underbrace{A\vec{v}_1 \downarrow \; \cdots}_{\substack{r \\ \neq \vec{0}}} \; \underbrace{\downarrow A\vec{v}_n}_{\substack{n-r \\ \vec{0}.}} \Big)$$

$\|A\vec{v}_i\| = \sqrt{\lambda_i}$

$$U \cdot \Sigma = \overset{U}{\begin{pmatrix} \downarrow & \cdots & \downarrow \\ u_1 & & u_m \end{pmatrix}} \overset{\Sigma}{\begin{pmatrix} \Sigma_{11} & & 0 & & \\ & \Sigma_{22} & & & 0 \\ 0 & & \Sigma_{33} & & \\ 0 & & & & 0 \end{pmatrix}}$$

$$= \begin{pmatrix} \Sigma_{11}\cdot\vec{u}_1 \downarrow & \cdots & \underset{\Sigma_{mm}\cdot u_m}{\downarrow} & \cdots 0 \end{pmatrix}$$

Construct U by define $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$.

Σ by define $\Sigma_{ii} = \sqrt{\lambda_i}$.

for $i \leq r$. when r: # of nonzero eigenvalues of $A^T \cdot A$.

Finally extend. $\{\vec{u}_1, \cdots, \vec{u}_r\}$ to an orthonormal basis of $\mathbb{R}^m$.

$\square$

Ex. $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}$ or $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix}$

$A^T \cdot A$ 2×2

$A^T \cdot A = 3\times 3$.

if $A = U \Sigma V^T$

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

pf. $B = A^T A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}$

$$= \begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix}$$

$$\det(B - \lambda I) = \begin{vmatrix} 6-\lambda & 12 \\ 12 & 24-\lambda \end{vmatrix} = \lambda^2 - 30\lambda + 24 \times 6 - 144$$

$$= \lambda^2 - 30\lambda + 0 = 0$$

$\Rightarrow \quad \lambda_2 = 0 \quad \lambda_1 = 30.$

For $\lambda_2 = 0$ $\begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$ $\quad \vec{x} = x_2 \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$

$$\vec{v}_2 = \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

For $\lambda_1 = 30.$ $\begin{pmatrix} -24 & 12 \\ 12 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$ $\quad \vec{x} = x_2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$\vec{v}_1 = \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

So. $V = \begin{pmatrix} \frac{1}{\sqrt{5}} 1 & \frac{1}{\sqrt{5}} 2 \\ \frac{1}{\sqrt{5}} 2 & \frac{1}{\sqrt{5}}(-1) \end{pmatrix}$ $\quad \Sigma = \underset{2}{{}_{3}\begin{pmatrix} \sqrt{30} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}$

$\vec{u}_1 = \dfrac{A\vec{v}_1}{\|A\vec{v}_1\|} \longrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} = \begin{pmatrix} 5 \\ 10 \\ 5 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} = \begin{pmatrix} \sqrt{5} \\ 2\sqrt{5} \\ \sqrt{5} \end{pmatrix}$

$\|A\vec{v}_i\| = \sqrt{\lambda_i}$

$$\vec{u}_1 = \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

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

$$\begin{pmatrix}1 & 2 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=0 \quad \Rightarrow \quad x_3 \ \ x_2 \text{ are fee}$$

$$x_1 = -x_3 - 2x_2$$

$$\vec{x} = \begin{pmatrix}-x_3-2x_2\\x_2\\x_3\end{pmatrix} = x_3\cdot\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

$$+ \ x_2\cdot\begin{pmatrix}-2\\1\\0\end{pmatrix}$$

So. $\vec{w}_2 = \begin{pmatrix}-1\\0\\1\end{pmatrix}$ $\quad \vec{w}_3 = \begin{pmatrix}-2\\1\\0\end{pmatrix}$

$\{\vec{u}_1, \vec{w}_2, \vec{w}_3\}$ is a set of basis for $\mathbb{R}^3$. and $\vec{u}_1 \perp \vec{w}_2, \vec{w}_3$.

Gram-Schmidt.

$$\vec{w}_3' = \vec{w}_3 - \frac{\vec{w}_3\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\cdot\vec{w}_2 = \begin{pmatrix}-2\\1\\0\end{pmatrix} - \frac{2}{2}\cdot\vec{w}_2$$

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

$\{\vec{u}_1, \overset{\vec{u}_2}{\overset{\shortparallel}{\frac{1}{\sqrt{2}}\vec{w}_2}}, \overset{\vec{u}_3}{\overset{\shortparallel}{\frac{\vec{w}_3'}{\|\vec{w}_3'\|}}}\}$ forms an orthonormal basis

$$U = \begin{pmatrix}1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3}\\ 2/\sqrt{6} & 0 & 1/\sqrt{3}\\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3}\end{pmatrix}.$$

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

Application: $rk(A)$ is really small. then.

$A \cdot x$ can be faster by writing $A = U \Sigma V^T$