03/13/2018

Recap

dim W and dim $W^\perp$

Let W be a subspace of $\mathbb{R}^n$. We have

$$\dim W + \dim W^\perp = n$$

Orthogonal set

Theorem 4, chapter 6

An orthogonal set of non-zero vectors is linearly independent.

Orthogonal basis

Suppose $\dim W = p$ and $\{\vec{v}_1, \cdots, \vec{v}_p\}$ is an orthogonal set of non-zero vectors in W.

$\Longrightarrow \{\vec{v}_1, \cdots, \vec{v}_p\}$ is an orthogonal basis for W.

Orthogonal projection

Theorem 8: Let $\{\vec{u}_1, \cdots, \vec{u}_p\}$ be an orthogonal basis for W (of $\mathbb{R}^n$).

Each $\vec{y}$ in $\mathbb{R}^n$ can be decomposed as

$$\vec{y} = \vec{\hat{y}} + \vec{z}$$

where $\vec{\hat{y}}$ is in W and $\vec{z}$ is in $W^\perp$

$$\vec{\hat{y}} = \frac{\vec{y}\cdot\vec{u_1}}{\vec{u_1}\cdot\vec{u_1}}\,\vec{u_1} + \cdots + \frac{\vec{y}\cdot\vec{u_p}}{\vec{u_p}\cdot\vec{u_p}}\,\vec{u_p}$$

Notation $\underbrace{\mathrm{proj}_W \vec{y}}_{\text{projection of } \vec{y} \text{ onto } W} \equiv \vec{\hat{y}}$

The decomposition is unique.

Theorem 9 (the best approximation theorem)

Let W be a subspace of $\mathbb{R}^n$

Let $\vec{y}$ be a vector in $\mathbb{R}^n$

We have $\|\vec{y} - \mathrm{proj}_W \vec{y}\| < \|\vec{y} - \vec{v}\|$

for all $\vec{v} \neq \mathrm{proj}_W \vec{y}$ in W.

(of all vectors in W, $\mathrm{proj}_W \vec{y}$ provides the best approximation to $\vec{y}$)

Def. The distance from a vector to a subspace

Let W be a subspace of $\mathbb{R}^n$

$$\underbrace{\operatorname{dist}(\vec{y}, W)}_{\text{distance from } \vec{y} \text{ to } W} \equiv \| \vec{y} - \operatorname{proj}_W \vec{y} \|$$

Sec. 6.5. Least squares problems.

Consider $A\vec{x} = \vec{b}$

A : $m \times n$

$\vec{x}$ is in $\mathbb{R}^n$

$\vec{b}$ is in $\mathbb{R}^m$

$A = [\vec{a}_1 \ \vec{a}_2 \cdots \vec{a}_n]$, $\quad \vec{a}_j$ is in $\mathbb{R}^m$

$A\vec{x} = \vec{b}$ may be inconsistent.

We can try to get $A\vec{x}$ as close to $\vec{b}$ as possible.

Def: A least-squares solution of $A\vec{x} = \vec{b}$ is an $\hat{\vec{x}}$ such that

$$\| \vec{b} - A\hat{\vec{x}} \| \leq \| \vec{b} - A\vec{x} \| \quad \text{for all } \vec{x} \text{ in } \mathbb{R}^n$$

We try to find $\displaystyle \min_{\vec{x} \text{ in } \mathbb{R}^n} \| \vec{b} - A\vec{x} \|$

Let $W = \text{Col } A$, a subspace of $\mathbb{R}^m$

$\vec{v}$ is in Col A

$\iff \vec{v} = A\vec{x}$ for some $\vec{x}$ in $\mathbb{R}^n$

We write

$$W = \{A\vec{x} \mid \vec{x} \text{ is in } \mathbb{R}^n\}$$

$$\min_{\vec{x} \text{ in } \mathbb{R}^n} \|\vec{b} - A\vec{x}\| = \min_{\vec{v} \text{ in } W} \|\vec{b} - \vec{v}\|$$

Let $\hat{\vec{b}} = \text{proj}_W \vec{b}$

$\Rightarrow \|\vec{b} - \hat{\vec{b}}\| \leq \|\vec{b} - \vec{v}\|$. for all $\vec{v}$ in W

Let $\hat{\vec{x}}$ be a solution of $A\hat{\vec{x}} = \hat{\vec{b}}$.

$$\|\vec{b} - A\hat{\vec{x}}\| \leq \|\vec{b} - A\vec{x}\| \text{ for all } \vec{x} \text{ in } \mathbb{R}^n$$

Conclusion:

*). A solution of $A\hat{\vec{x}} = \hat{\vec{b}}$ is a least-squares solution of $A\vec{x} = \vec{b}$

*) Conversely, a least-squares solution of $A\vec{x} = \vec{b}$ satisfies $A\hat{\vec{x}} = \hat{\vec{b}}$

Statement:

$$\left(\text{The set of least-squares solutions of } A\vec{x}=\vec{b}\right)$$

$$= \left(\text{the solution set of } A\hat{\vec{x}}=\hat{\vec{b}}\right)$$

$$\hat{\vec{b}} = \text{proj}_W \vec{b} \qquad W = \text{Col} A$$

Existence of solution:

$$\hat{\vec{b}} = \text{proj}_W \vec{b} \quad \text{is in } W = \text{Col} A$$

$\Rightarrow \hat{\vec{b}}$ is a linear combination of columns of A

$\Rightarrow A\hat{\vec{x}} = \hat{\vec{b}}$ is consistent.

A method for calculating $\hat{\vec{x}}$:

$\text{proj}_W \vec{b}$ is difficult to calculate.

because we are not given an orthogonal basis for $W = \text{Col} A$.

Solving $A\hat{\vec{x}} = \hat{\vec{b}}$ is not practical.

We need a practical method.

Orthogonal decomposition of $\vec{b}$

$$\vec{b} = \underbrace{\hat{\vec{b}}}_{\text{in } W} + \underbrace{\vec{z}}_{\text{in } W^{\perp}}$$

⑥

$(\vec{b} - \hat{\vec{b}})$ is in $W^{\perp} = (\mathrm{Col}\, A)^{\perp}$

Recall $(\mathrm{Col}\, A)^{\perp} = \mathrm{Nul}\ A^T$

$\Longrightarrow \quad A^T(\vec{b} - \hat{\vec{b}}) = \vec{0}$

Conclusion *) A solution of $A\hat{\vec{x}} = \hat{\vec{b}}$ satisfies

$$A^T(\vec{b} - A\hat{\vec{x}}) = \vec{0}$$

We write it as

$$A^T\vec{b} - (A^TA)\hat{\vec{x}} = 0.$$

$$\Longrightarrow \quad (A^TA)\hat{\vec{x}} = A^T\vec{b}$$

This is called the normal equation for $A\vec{x} = \vec{b}$

*). Conversely, a solution of $(A^TA)\hat{\vec{x}} = A^T\vec{b}$ satisfies.

$$A\hat{\vec{x}} = \hat{\vec{b}}.$$

proof: $\quad A^T(\vec{b} - A\hat{\vec{x}}) = \vec{0}.$

Recall $\mathrm{Nul}\ A^T = (\mathrm{Col}\, A)^{\perp}$

$\Longrightarrow \quad (\vec{b} - \underset{\text{it is in } W}{A\hat{\vec{x}}})$ is in $(\mathrm{Col}\, A)^{\perp} = W^{\perp}$

$\Longrightarrow \quad A\hat{\vec{x}} = \mathrm{proj}_W \vec{b} = \hat{\vec{b}}$

Theorem 13

(The set of least-squares solutions of $A\vec{x}=\vec{b}$)
$= ($the solution set of $(A^TA)\vec{\hat{x}} = A^T\vec{b}$ $)$

Theorem 14

$A = m \times n$

The statements below are equivalent.

a) $A\vec{x}=\vec{b}$ has a unique least-squares solution.

b) The columns of A are linearly independent.

c) $\underline{A^TA}$ is invertible.

$$\underset{n\times m}{A^T} \cdot \underset{m\times n}{A} = \underset{n\times n}{[\quad]}$$

We show a) $\Longleftrightarrow$ b).

$A\vec{\hat{x}}=\vec{b}$ has a unique solution.

$\Longleftrightarrow$ $A\vec{\hat{x}}=\vec{0}$ has only the trivial solution.

$\Longleftrightarrow$ Columns of A are linearly independent.
$\leftarrow$ b).

Next we show a) $\Longleftrightarrow$ c).

a) $\Longleftrightarrow$ $(A^TA)\vec{\hat{x}} = A^T\vec{b}$ has a unique solution.

$\Longleftrightarrow$ $(A^TA)\vec{\hat{x}}=\vec{0}$ has only the trivial solution.

$\iff (A^TA)$ is invertible. $\Leftarrow$ c).

Summary: $A\vec{x}=\vec{b}$ always has a least-squares solution.

It has a unique least-squares solution if (A^TA) is invertible.

Sec. 7.1 Diagonalization of symmetric matrices.

Def: If $A^T=A$, we say A is symmetric

Ex. $A=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}$ is symmetric

$A^T=A$ if and only if $a_{ij}=a_{ji}$ for all i, j

$A=\begin{bmatrix}1 & 2\\ 3 & 3\end{bmatrix}$ is not symmetric

Statement: If A is symmetric, then all eigenvalues are real.

Key step: Complex inner product

$$\vec{u}\cdot\vec{v} = \vec{u}^T\cdot\overline{\vec{v}} = u_1\overline{v_1}+u_2\overline{v_2}+\cdots+u_n\overline{v_n}$$

$\vdots$

Theorem 1:

If A is symmetric, then eigenvectors for different eigenvalues are orthogonal to each other.

proof:

$$A\vec{V}_1 = \lambda_1 \vec{V}_1$$

$$A\vec{V}_2 = \lambda_2 \vec{V}_2, \qquad \lambda_1 \neq \lambda_2$$

$$\begin{aligned}
\underbrace{(A\vec{V}_1)\cdot\vec{V}_2}_{=\,\lambda_1 \vec{V}_1\cdot\vec{V}_2} &= (A\vec{V}_1)^T\vec{V}_2 = (\vec{V}_1^{\,T}A^T)\vec{V}_2 \\
&= \vec{V}_1^{\,T}(A^T\vec{V}_2) \\
&= \vec{V}_1^{\,T}(A\vec{V}_2) \\
&= \vec{V}_1\cdot(A\vec{V}_2) \\
&= \lambda_2 \vec{V}_1\cdot\vec{V}_2
\end{aligned}$$

$$\Longrightarrow \lambda_1 \vec{V}_1\cdot\vec{V}_2 = \lambda_2 \vec{V}_1\cdot\vec{V}_2$$

$$\Longrightarrow (\lambda_1-\lambda_2)\,\vec{V}_1\cdot\vec{V}_2 = 0$$

$$\Longrightarrow \vec{V}_1\cdot\vec{V}_2 = 0$$

Special case:

Suppose an $n\times n$ symmetric matrix A has n distinct eigenvalues

$$\lambda_1, \lambda_2, \cdots, \lambda_n.$$

Eigenvectors: $\vec{V}_1, \vec{V}_2, \cdots, \vec{V}_n$

$\Rightarrow \{\vec{v}_1, \cdots, \vec{v}_n\}$ is an orthogonal set.

Def: An orthonormal set is an orthogonal set of unit vectors.

We can make $\{\vec{v}_1, \cdots, \vec{v}_n\}$ an orthonormal set.

Let $P = [\vec{v}_1 \cdots \vec{v}_n]$. $P: n \times n$.

$$P^T P = \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} [\vec{v}_1 \cdots \vec{v}_n] = \begin{bmatrix} \vec{v}_1^T \vec{v}_1 & \vec{v}_1^T \vec{v}_2 & \cdots & \vec{v}_1^T \vec{v}_n \\ \vdots & & & \\ \vec{v}_n^T \vec{v}_1 & \cdots & \cdots & \vec{v}_n^T \vec{v}_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

Def: If $P^{-1} = P^T$, we say P is an orthogonal matrix

Def: We say A is orthogonally diagonalizable if there exists an orthogonal matrix P such that

$$P^T A P = D$$

Statement:

Suppose an $n \times n$ symmetric matrix A has n distinct eigenvalues.

Then A is orthogonally diagonalizable.

Theorem 2:

Every symmetric matrix is orthogonally diagonalizable.