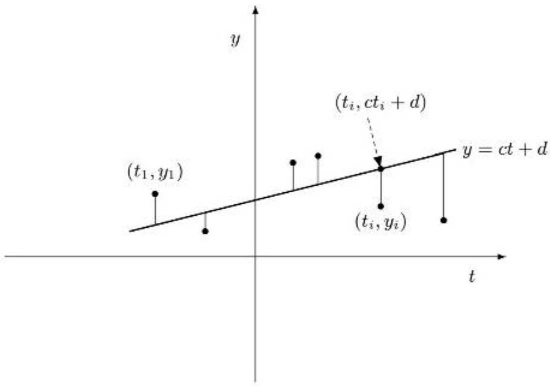
Linear Algebra Class on 1 June

Seanie Lee

1 June 2019

6.3 Least Square



We want to find c, d such that minimize $\sum_{i=1}^{k}(y_i - (ct_i + d))^2$.

$A = \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_k & 1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} c \\ d \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}$, then it is equivalent to find $\mathbf{x}$ such that minimizes $\|\mathbf{y} - A\mathbf{x}\|^2$.

In other words, we want to find $\mathbf{x}_0 = (c_0, d_0)$ such that $\|\mathbf{y} - A\mathbf{x}_0\| \leq \|\mathbf{y} - A\mathbf{x}\|$ for all $\mathbf{x} \in \boldsymbol{V}$. First we show the existence of $\mathbf{x}_0$ by corollary to the theorem 6.6.

Corollary. $\|\mathbf{v} - \mathbf{w}\| \leq \|\mathbf{v} - \mathbf{x}\|$ *for all* $\mathbf{x} \in \boldsymbol{W} \leq \boldsymbol{V}$

Proof. $\mathbf{v} := \mathbf{w} + \mathbf{w}' \quad \mathbf{w} \in \boldsymbol{W}, \mathbf{w}' \in \boldsymbol{W}^{\perp}$

$$\begin{aligned}
\|\mathbf{v} - \mathbf{x}\|^2 &= \|\mathbf{w} + \mathbf{w}' - \mathbf{x}\|^2 \\
&= \|(\mathbf{w} - \mathbf{x}) + \mathbf{w}'\|^2 \\
&= \langle (\mathbf{w} - \mathbf{x}) + \mathbf{w}', (\mathbf{w} - \mathbf{x}) + \mathbf{w}' \rangle \\
&= \langle \mathbf{w} - \mathbf{x}, \mathbf{w} - \mathbf{x} \rangle + \langle \mathbf{w} - \mathbf{x}, \mathbf{w}' \rangle + \langle \mathbf{w}', \mathbf{w} - \mathbf{x} \rangle + \langle \mathbf{w}', \mathbf{w}' \rangle \\
&= \|\mathbf{w} - \mathbf{x}\|^2 + \|\mathbf{w}'\|^2 \\
&\geq \|\mathbf{w}'\|^2 = \|\mathbf{v} - \mathbf{w}\|^2
\end{aligned}$$

Since $\|\cdot\| \geq 0$, $\|\mathbf{v} - \mathbf{x}\| \geq \|\mathbf{v} - \mathbf{w}\|$ $\square$

By taking $\boldsymbol{W} := \{A\mathbf{x} | \mathbf{x} \in \mathbb{C}^n\}$, we can find such $\mathbf{x}_0 \in \mathbb{C}^n$. Then is such an $\mathbf{x}_0$ is unique?

$$\begin{aligned} L_A \text{ is 1-1} &\iff kerL_A = \mathbf{O} \\ &\iff \dim\, imL_A = n \\ &\iff \dim\, \mathcal{C} = n \\ &\iff \text{rankA} = n \end{aligned}$$

Lemma 6.1. $\langle A\mathbf{x},\, \mathbf{y}\rangle_m = \langle \mathbf{x},\, A^*\mathbf{y}\rangle_n$ *where* $A \in M_{m\times n}(\mathbb{C})$, $\mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m$, *and* $\langle \cdot, \cdot\rangle$ *is dot product.*

Proof.

$$\begin{aligned} \langle A\mathbf{x},\, \mathbf{y}\rangle_m &= \mathbf{y}^*(A\mathbf{x}) \\ &= \overline{(A\mathbf{x})^*\mathbf{y}} \\ (\because \langle A\mathbf{x},\, \mathbf{y}\rangle &= \overline{\langle \mathbf{y},\, A\mathbf{x}\rangle}\ ) \\ &= \overline{\mathbf{x}^*(A^*\mathbf{y})} \\ &= \overline{\langle A^*\mathbf{y},\, \mathbf{x}\rangle} \\ &= \langle \mathbf{x},\, A^*\mathbf{y}\rangle_n \end{aligned}$$

□

Lemma 6.2. *rankA* $=$ *rank(*A^*A*)* $A \in \mathfrak{M}_{m\times n}(\mathbb{C})$

Proof. By the Dimension Theorem, $\dim(imL_A) + \dim(kerL_A) = n$. Thus it suffices to show that $A\mathbf{x} = \mathbf{0} \iff A^*A = \mathbf{0}$. $\Longrightarrow$ it is trivial.
$\Longleftarrow$

$$\begin{aligned} \langle A\mathbf{x},\, A\mathbf{x}\rangle &= \langle \mathbf{x},\, A^*A\mathbf{x}\rangle \quad (\because \text{ by the Lemma 6.1}) \\ &= \langle x,\, \mathbf{0}\rangle \\ &= 0 \\ \therefore A\mathbf{x} &= \mathbf{0} \end{aligned}$$

□

By the Theorem 6.6 and its corollary, $A\mathbf{x}_0 - \mathbf{y} \in \boldsymbol{W}^\perp$ and $\forall \mathbf{x}, \langle A\mathbf{x},\, A\mathbf{x}_0 - \mathbf{y}\rangle$. By the Lemma 6.1 $\langle A\mathbf{x},\, A\mathbf{x}_0 - \mathbf{y}\rangle = \langle \mathbf{x},\, A^*(A\mathbf{x}_0 - \mathbf{y})\rangle = 0$. Thus $A^*A\mathbf{x}_0 - A^*\mathbf{y} = \mathbf{0}$. If rank$A = n$(full rank), then by the Lemma 6.2, A^*A is invertible. $\therefore \mathbf{x}_0 = (A^*A)^{-1}A^*\mathbf{y}$

6.4 Normal and Adjoint Operator

Theorem. *If* $\boldsymbol{T}\mathbf{v} = \lambda\mathbf{v} \Longrightarrow \boldsymbol{T}^*$ *has eigenvector*

Proof.

$$
\begin{aligned}
(\boldsymbol{T}-\lambda\boldsymbol{I})\mathbf{v} &= \mathbf{0} \\
\forall\mathbf{x} \quad 0 = \langle\mathbf{0},\, \mathbf{x}\rangle &= \langle(\boldsymbol{T}-\lambda\boldsymbol{I})\mathbf{v},\, \mathbf{x}\rangle \\
&= \langle\mathbf{v},\, (\boldsymbol{T}-\lambda\boldsymbol{I})^*\mathbf{x}\rangle \\
&= \langle\mathbf{v},\, (\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I})\mathbf{x}\rangle \\
(\because [(\boldsymbol{T}-\lambda\boldsymbol{I})^*]_\beta &= [\boldsymbol{T}-\lambda\boldsymbol{I}]_\beta^* \\
&= [\boldsymbol{T}]_\beta^* - [\lambda\boldsymbol{I}]_\beta^* \\
&= [\boldsymbol{T}^*]_\beta - \overline{\lambda}[\boldsymbol{I}]_\beta \\
\therefore (\boldsymbol{T}-\lambda\boldsymbol{I})^* &= \boldsymbol{T}^* - \overline{\lambda}\boldsymbol{I} \quad)
\end{aligned}
$$

We want to show that $ker(\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I}) \neq \mathbf{O}$. Since $\mathbf{v}$ is orthogonal to $(\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I})\mathbf{v} \quad \forall\mathbf{x}$, $\mathbf{v} \perp im(\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I})$. Since $\mathbf{v} \neq \mathbf{0}$, $(im(\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I}))^\perp \neq \mathbf{O}$, so that $im(\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I}) \nleq \mathbb{C}^n$. Thus $\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I}$ is not onto and by the pigeonhole's principle $\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I}$ is not 1-1. Thus $\exists\mathbf{v}_0 \neq \mathbf{0}$ such that $(\boldsymbol{T}^*-\overline{\lambda}\boldsymbol{I})\mathbf{v}_0 = \mathbf{0}$.
$\therefore \boldsymbol{T}^*\mathbf{v}_0 = \overline{\lambda}\mathbf{v}_0$ ☐

Theorem (Schur)**.** *If $\phi_T(t)$ splits, then there is an orthonormal basis β such that $[\boldsymbol{T}]_\beta$ is upper-triangular.*

Proof. It's up to you :) ☐

Assume that there is an orthogonal basis consisting of eigenvectors of $\boldsymbol{T}$. Then $[\boldsymbol{T}]_\beta$ is diagonal. Since $[\boldsymbol{T}^*]_\beta = [\boldsymbol{T}]_\beta^*$, $[\boldsymbol{T}^*]_\beta$ is also diagonal. Since diagonal matrices commute, so $\boldsymbol{T}$ and $\boldsymbol{T}^*$ commute. I.e., $\boldsymbol{T}\boldsymbol{T}^* = \boldsymbol{T}^*\boldsymbol{T}$

Definition 6.1. *$\boldsymbol{T}$ is normal if $\boldsymbol{T}\boldsymbol{T}^* = \boldsymbol{T}^*\boldsymbol{T}$*

Note. *Normal operator $\boldsymbol{T}$ does not guarantee the existence of eigenvectors.*
$A = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$ *is rotation matrix $(0 < \theta < \pi)$. A has no eigenvectors, but A is normal.*

Theorem.

(a) $\|\boldsymbol{T}\mathbf{x}\| = \|\boldsymbol{T}^*\mathbf{x}\| \ \forall\mathbf{x} \in \boldsymbol{V}$

(b) $\boldsymbol{T} - c\boldsymbol{I}$ *is normal* $\forall c \in \mathbb{C}$

(c) $\boldsymbol{T}\mathbf{x} = \mathbf{x} \Longrightarrow \boldsymbol{T}^* = \overline{\lambda}\mathbf{x}$

(d) $\lambda_1 = \lambda_2$ *are eignvalues of* $\boldsymbol{T}$, $\mathbf{x}_1, \mathbf{x}_2$ *are corresponding eigenvectors of* $\boldsymbol{T} \Longrightarrow \mathbf{x}_1 \perp \mathbf{x}_2$

Proof.

(a)

$$
\begin{aligned}
\|\boldsymbol{T}\mathbf{x}\|^2 &= \langle\boldsymbol{T}\mathbf{x},\, \boldsymbol{T}\mathbf{x}\rangle \\
&= \langle\mathbf{x},\, \boldsymbol{T}^*\boldsymbol{T}\mathbf{v}\rangle \\
&= \langle\mathbf{x},\, \boldsymbol{T}\boldsymbol{T}^*\mathbf{x}\rangle \\
&= \langle\boldsymbol{T}^*\mathbf{x},\, \boldsymbol{T}^*\mathbf{x}\rangle
\end{aligned}
$$

$(\because \boldsymbol{T}^*$ is linear such that $\langle \boldsymbol{T}\mathbf{x},\, \mathbf{x}\rangle = \langle \mathbf{x},\, \boldsymbol{T}^*\mathbf{x}\rangle$. $\langle \mathbf{x},\, \boldsymbol{T}\mathbf{x}\rangle = \overline{\langle \boldsymbol{T}\mathbf{x},\, \mathbf{x}\rangle} = \overline{\langle \mathbf{x},\, \boldsymbol{T}^*\mathbf{x}\rangle} = \langle \boldsymbol{T}^*\mathbf{x},\, \mathbf{x}\rangle$)

(b)

$$\begin{aligned}
(\boldsymbol{T} - c\boldsymbol{I})^*(\boldsymbol{T} - c\boldsymbol{I}) &= (\boldsymbol{T}^* - \bar{c}\boldsymbol{I})(\boldsymbol{T} - c\boldsymbol{I}) \\
&= \boldsymbol{T}^*\boldsymbol{T} - \bar{c}\boldsymbol{T} - c\boldsymbol{T}^* - \bar{c}c\boldsymbol{I} \\
&= \boldsymbol{T}\boldsymbol{T}^* - \bar{c}\boldsymbol{T} - c\boldsymbol{T}^* - \bar{c}c\boldsymbol{I} \\
&= (\boldsymbol{T} - c\boldsymbol{I})(\boldsymbol{T}^* - \bar{c}\boldsymbol{I}) \\
&= (\boldsymbol{T} - c\boldsymbol{I})(\boldsymbol{T} - c\boldsymbol{I})^*
\end{aligned}$$

(c)

$$\begin{aligned}
(\boldsymbol{T} - \lambda\boldsymbol{I})\mathbf{x} &= \mathbf{0} \\
0 = \|(\boldsymbol{T} - \lambda\boldsymbol{I})\mathbf{x}\| &= \left\|(\boldsymbol{T}^* - \bar{\lambda}\boldsymbol{I})\mathbf{x}\right\| \text{ (by (a))} \\
(\boldsymbol{T}^* - \bar{\lambda}\boldsymbol{I})\mathbf{x} &= \mathbf{0} \\
\therefore \boldsymbol{T}^*\mathbf{x} &= \bar{\lambda}\mathbf{x}
\end{aligned}$$

(d)

$$\begin{aligned}
\lambda_1\langle \mathbf{x}_1,\, \mathbf{x}_2\rangle &= \langle \lambda_1\mathbf{x}_1,\, \mathbf{x}_2\rangle \\
&= \langle \boldsymbol{T}\mathbf{x}_1,\, \mathbf{x}_2\rangle \\
&= \langle \mathbf{x}_1,\, \boldsymbol{T}^*\mathbf{x}_2\rangle \\
&= \langle \mathbf{x}_1,\, \overline{\lambda_2}\mathbf{x}_2\rangle \\
&= \lambda_2\langle \mathbf{x}_1,\, \mathbf{x}_2\rangle \\
(\lambda_1 - \lambda_2)\langle \mathbf{x}_1,\, \mathbf{x}_2\rangle &= 0 \\
\text{Since } \lambda_1 \neq \lambda_2, \langle \mathbf{x}_1,\, \mathbf{x}_2\rangle &= 0. \therefore \mathbf{x}_1 \perp \mathbf{x}_2
\end{aligned}$$

$\square$