CPSC 531F March 25, 2021

- Finish remarks on variation methods
- Perron-Frobenius thm on Tuesday
- Then more applications

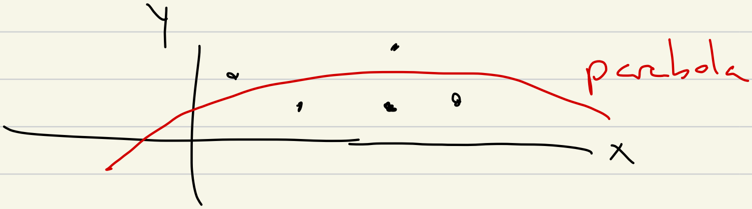
Simpler Variational Method:

Least squares fit:

Idea from applications:

you have n data points,

$(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n) \in \mathbb{R}^2$



Model: theoretically:

$$y = f(x) = a_1 \cdot 1 + a_2 x + a_3 x^2 + a_4 e^{-x^2}/\sqrt{x}$$

$$= a_1 f_1(x) + \dots + a_m f_m(x)$$

Error:

Error $(a_1, \dots, a_m)$

$$:= \sum_i |y_i - \underbrace{f(x_i)}_{a_1 f_1(x_i) + \dots + a_m f_m(x_i)}|^2$$

$$= \left\| \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} f_1(x_1) \\ \vdots \\ f_1(x_n) \end{bmatrix} a_1 - \dots - \begin{bmatrix} f_m(x_1) \\ \vdots \\ f_m(x_n) \end{bmatrix} a_m \right\|_2^2$$

Abstractly!

$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \vec{f}_1 = \begin{bmatrix} f_1(x_1) \\ \vdots \\ f_1(x_n) \end{bmatrix},$$

$$\vec{f}_2, \ldots, \vec{f}_m$$

Searching for $a_1, \ldots, a_m$ s.t. min value of

$$\left\| \vec{y} - \vec{f}_1 a_1 - \cdots - \vec{f}_m a_m \right\|_2^2$$

$\vec{y}, \vec{f}_1, \ldots, \vec{f}_m \in \mathbb{R}^n$

Variational Method!

① Imagine that we have

$a_1, \dots, a_m$ where min occurs.

$$\mathcal{E}(a_1, \dots, a_m) = \left\| \vec{y} - \begin{bmatrix} | & | & \\ f_1 & f_2 & \dots \\ | & | & \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \right\|_2^2$$

$$= \| \underbrace{\vec{y}}_{\text{given}} - \underbrace{F}_{\text{matrix}} \vec{a} \|_2^2$$

Take any $\vec{b} \in \mathbb{R}^m$ consider

$$g_{\vec{b}}(\varepsilon) = \| \vec{y} - F(\vec{a} + \varepsilon \vec{b}) \|_2$$

Err is local min

$$g(\varepsilon) = c_0 + c_1\varepsilon + \underbrace{\text{higher order}}_{c_2\varepsilon^2}$$

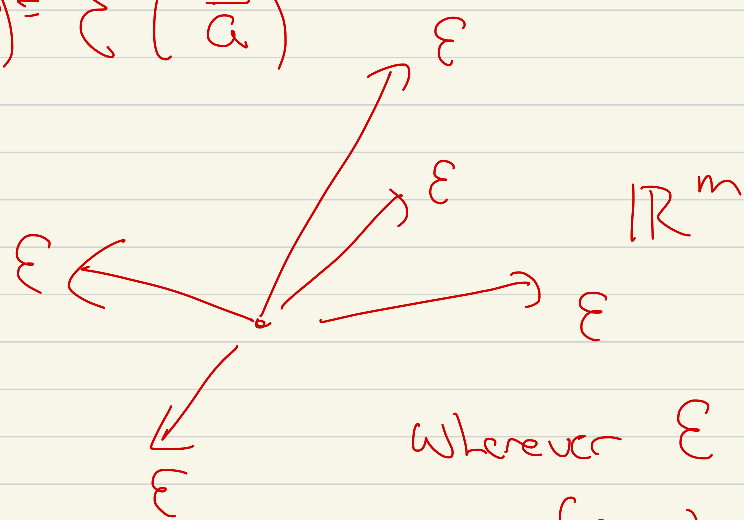
$$\shortparallel$$

$$g_{\vec{b}}(\varepsilon)$$

$$c_1 = 0$$

$$=$$

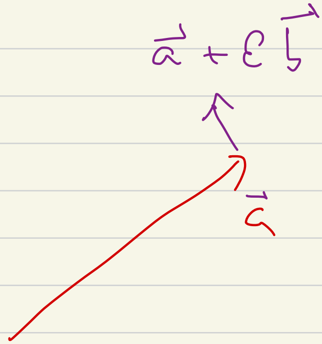
$$\text{Error}(\vec{a}) = \mathcal{E}(\vec{a})$$



Wherever $\mathcal{E}$ = local min (max), $\vec{a}$

$$\nabla\mathcal{E}(\vec{a}) = 0$$

$\mathcal{E}_{rr}$:



$\mathcal{E}_{rr}$ has local min/max

or just any critical point,

$\nabla \mathcal{E}_{rr} = 0$

$$g(\varepsilon) = g_{\vec{b}}(\varepsilon)$$

$$= \text{Error}(\vec{a} + \varepsilon \vec{b})$$

$$g'(0) = 0$$

Calculation!

$$\text{Error}(\vec{a} + \varepsilon \vec{b})$$

$$= \left(\vec{y} - F(\vec{a} + \varepsilon \vec{b}), \vec{y} - F(\vec{a} + \varepsilon \vec{b}) \right)$$

$$= C_0 + \varepsilon C_1 + \varepsilon^2 C_2$$

ε part of

$$\left((\vec{y} - F\vec{a}) - \varepsilon F\vec{b},\ (\vec{y} - F\vec{a}) - \varepsilon F\vec{b} \right)$$

ε part: $\quad -2\left(\vec{y} - F\vec{a},\ F\vec{b} \right)$

dot product $\quad (\vec{x} \cdot A\vec{y}) = (A^T\vec{x}) \cdot \vec{y}$

or $\quad (\vec{x}, \mathcal{L}\vec{y}) = (\mathcal{L}^*\vec{x}, \vec{y})$ more generally

If

$$(\vec{y} - F\vec{a}, F\vec{b}) = 0$$

for all $\vec{b}$, then

$$\Downarrow$$

for all $\vec{b}$ $\quad (F^T(\vec{y} - F\vec{a}), \vec{b}) = 0$

$$\Rightarrow \quad F^T(\vec{y} - F\vec{a}) = 0$$

$$= \quad \boxed{F^T F \vec{a} = F^T \vec{y}}$$

normal equations

Example:

$$y = a_1 + a_2 x + a_3 x^2$$

$$\min \left\| \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{bmatrix} - a_1 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - a_2 \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} - a_3 \begin{bmatrix} x_1^2 \\ \vdots \\ \vdots \\ x_n^2 \end{bmatrix} \right\|$$

could replace by $\begin{bmatrix} e^{-x_1} \\ \vdots \\ e^{-x_n} \end{bmatrix}$

$$\vec{y} - F \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$F = \begin{array}{ccc} \vec{f_1} & \vec{f_2} & \vec{f_3} \\ \downarrow & \downarrow & \downarrow \end{array}$$

$$F = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}$$

$$F^T F \vec{a} = F^T \vec{y} \quad \text{here}$$

$$\begin{bmatrix} - f_1^T - \\ - f_2^T - \\ - f_3^T - \end{bmatrix} \begin{bmatrix} | & | & | \\ f_1 & f_2 & f_3 \\ | & | & | \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} - f_1^T - \\ - f_2^T - \\ - f_3^T - \end{pmatrix} \vec{y}$$

$$\begin{bmatrix} f_1 \cdot f_1 & f_1 \cdot f_2 & f_1 \cdot f_3 \\ - & \cdot & \cdot \\ - & \cdot & \cdot \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{bmatrix} f_1 \cdot \vec{y} \\ f_2 \cdot \vec{y} \\ f_3 \cdot \vec{y} \end{bmatrix}$$

$=$

even $\quad y = f(x, w, z)$

data $\quad y_1, x_1, w_1, z_1, \ldots$

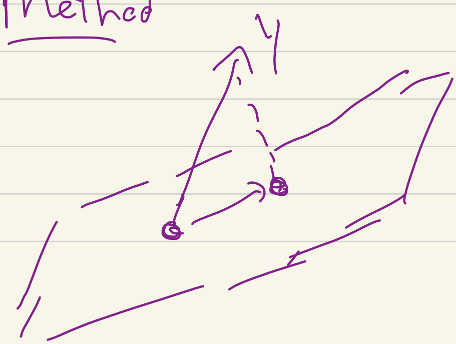
e.g.,
$$f_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad f_2 = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad f_3 = \begin{bmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix},$$

normal eqs.

Least squares fit

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \end{bmatrix}$$

Variational Method



Subspace spanned by $\vec{f}_1, \dots, \vec{f}_m$

Fit to a line $y = a_1 + a_2 x$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \end{bmatrix}$$

= Cald const

$$\text{Error}_{L^p} := \| \vec{y} - a_1 \vec{f}_1 - \dots + a_m \vec{f}_m \|_p .$$

=

Least squares! say you want to add

$$(x_{n+1}, y_{n+1})$$

If you keep values

keep running tally

$n, \sum x_i, \sum x_i^2$ (other sums like $\sum x_i^3, \sum x_i^4$)

$\sum y_i, \sum y_i x_i$ (and for parabolic fit $\sum y_i x_i^2$)

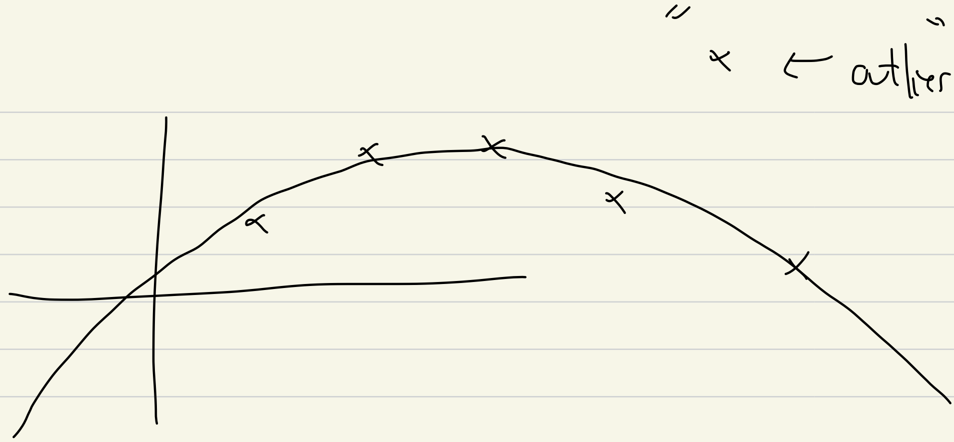
To solve best least squares fit

$y = a_1 + a_2 x, \quad y = a_1 + a_2 x + a_3 x^2, \ldots$

You ~~can~~ easily add point

(x_{n+1}, y_{n+1})

or delete points (x_i, y_i)



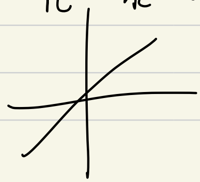
Rayleigh quotient: $A \in M_n(\mathbb{R})$, symmetric

$$R_A = \frac{(A\vec{v}, \vec{v})}{(\vec{v}, \vec{v})}$$

assume R_A has maximum at $\vec{v}$,

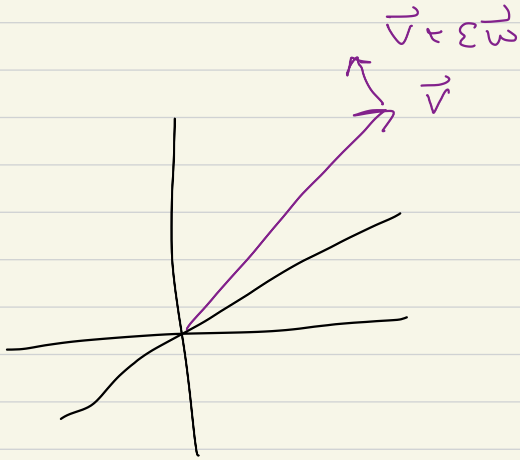
consider $\nabla R_A = 0$

$R : \mathbb{R}^n \to \mathbb{R}$



If so

$$g(\varepsilon) = R_A(\vec{v} + \varepsilon \vec{w})$$



If $\nabla R_A = 0$ at $\vec{v}$, then

$$A\vec{v} = \lambda \vec{v} \text{ for some } \lambda.$$

e.g. $A = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$

$$R_A(\vec{v}) = \frac{v_1^2 d_1 + v_2^2 d_2 + v_3^2 d_3}{v_1^2 + v_2^2 + v_3^2}$$

So R_A has $\left\{\begin{array}{l} \text{max} \\ \text{min} \\ \text{"saddle"} \end{array}\right.$ at $\vec{v}$

really!

$$(\nabla R_A)(\vec{v}) = 0$$

So $\vec{v}$ is a $\left\{\begin{array}{l} \text{critical point} \\ \text{stationary value} \end{array}\right\}$ of R_A

then $\qquad A\vec{v} = \lambda\vec{v}$.

But

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

then $\qquad A\vec{e}_1 = 3\vec{e}_1, \quad A\vec{e}_2 = 2\vec{e}_2, \quad A\vec{e}_3 = \vec{e}_3$

and $\quad \underline{\vec{e}_1}, \underline{\vec{e}_2}, \underline{\vec{e}_3}$ are all

$\left\{ \begin{matrix} \text{stationary values} \\ \text{critical points} \end{matrix} \right\}$ of R_A

Recall!

$$A = \begin{bmatrix} 3 & & \\ & 2 & \\ & & 1 \end{bmatrix},$$

$\vec{e}_1$ local (global) max, $R_A = 3$

$\vec{e}_2$ "saddle point" $\nabla R_A = 0$

$\vec{e}_3$ local (global) min, $R_A = 1$

=

(Similar to SVD)

=

Variation Method !

- Least Squares single crit point
- Rayleigh quotient
- SVD

$\Big\}$ typically n crit pts

Rayleigh quotient: at any critical point $\vec{v}$

$$A\vec{v} = \lambda\vec{v}$$

and $$\lambda = R_A(\vec{v})$$

So maximize $R_A \Rightarrow \lambda$ largest eig

minimize $R_A \Rightarrow$ " min "

SVD: Given $A \in M_{m,n}(\mathbb{R})$

$$f(\vec{u},\vec{v}) = \min \left\| A - \vec{u}\,\vec{v}^T \right\|_{frob}$$

for $\vec{u},\vec{v}$ minimize,

$$\text{(1)} \quad \begin{cases} A^T A \vec{v} = \lambda \vec{v} \\ A A^T \vec{u} = \lambda \vec{u} \end{cases} \quad \text{same } \lambda$$

$$\left. \begin{aligned} A\vec{v} &= \alpha \vec{u} \\ A^T \vec{u} &= \beta \vec{v} \end{aligned} \right\} \quad \begin{aligned} \alpha &= \vec{v}\cdot\vec{v} \\ \beta &= \vec{u}\cdot\vec{u} \end{aligned}$$

From variational principle...

Break 4 min.

After break!

n large

want to reduce n

$\vec{x}_1, \dots, \vec{x}_m \in \mathbb{R}^n$ ← given

$\downarrow$

$\vec{z}_1, \dots, \vec{z}_m \in \mathbb{R}^k$ ← looking

k small

take $k=1$

Best $\vec{z}_1, \dots, \vec{z}_m \in \mathbb{R}^k$, $\mathcal{L} : \mathbb{R}^k \to \mathbb{R}^n$

s.t.

$$f(\vec{z}_1, \dots, \vec{z}_m, \mathcal{L}) = \sum_i \|\vec{x}_i - \mathcal{L}\vec{z}_i\|^2$$

as small as possible

Rem: Consider

$$f(\vec{u},\vec{v}) = \left\| A - \vec{u}\,\vec{v}^T \right\|^2_{\text{Frob}}$$

Whenever $\vec{v}$ eigenvector of A^TA

$$A^T A \vec{v} = \lambda \vec{v}, \quad \text{take } \lambda > 0$$

and

$$\vec{u} = \frac{A\vec{v}}{\vec{v}\cdot\vec{v}}$$

then

$(\vec{u},\vec{v})$ is a critical point for f

in each variable

$k = 1$

given $\vec{x}_1, \ldots, \vec{x}_m \in \mathbb{R}^n$

$$\begin{bmatrix} - & \vec{x}_1^T & - \\ & \vdots & \\ - & \vec{x}_m^T & - \end{bmatrix} = X \in M_{m,n}(\mathbb{R})$$

looking for $z_1, \ldots, z_m \in \mathbb{R}$

$\mathcal{L} : \mathbb{R} \to \mathbb{R}^n$

i.e. $1 \mapsto \vec{y} \in \mathbb{R}^n$

then

$$\alpha \mapsto \alpha \vec{y} \in \mathbb{R}^n$$

s.t.

$$\left\| X - \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix} y^T \right\|_{Frob}$$

$\Downarrow$

$$\left\| \begin{bmatrix} - & x_1^T & - \\ & & \\ - & x_m^T & - \end{bmatrix} - \begin{bmatrix} - & z_1\, y^T & - \\ - & z_2\, y^T & - \\ & \vdots & \\ - & z_m\, y^T & - \end{bmatrix} \right\|_{Frob}$$

minimized

$$\| X - \vec{z}\, \vec{y}^T \| \quad \text{minimized}$$

we know min:

$\vec{z} = \vec{u}$, $\vec{y} = \vec{v}$ s.t. above eqs hold.

So! there exist $\vec{z}, \vec{y}$ minimizing this $\| \; \|_{Frob}$, SVD,

$$\boxed{\begin{bmatrix} — & x_1^T & — \\ & \vdots & \\ — & x_n^T & — \end{bmatrix} \vec{y} = \alpha \vec{z} = \alpha \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}}$$

So

$$z_i = \frac{\vec{x}_i \cdot \vec{y}}{\alpha}$$

So z_i is linear func of $\vec{x}_i$

dimension reduction, linear!!

Class Ends

E.g.,

$$f(x_1, x_2, x_3) = \frac{5x_1^2 + 4x_2^2 + 3x_3^2}{x_1^2 + x_2^2 + x_3^2}$$

Claim:

$$\nabla f = 0 \iff \vec{x} \text{ either } \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1} = f_{x_1} = \frac{\partial}{\partial x_1}\left(\frac{\text{numerator}}{\text{denominator}}\right)$$

$$= \frac{(\text{denom})(\text{num})_{x_1} - (\text{denom})_{x_1}(\text{num})}{(\text{denom})^2}$$

$$= \frac{(x_1^2+x_2^2+x_3^2)\,10x_1 - (2x_1)(5x_1^2+4x_2^2+3x_3^2)}{(x_1^2+x_2^2+x_3^2)^2}$$

$$= \underline{10x_1(x_2^2+x_3^2) - 2x_1(4x_2^2+3x_3^2)}$$

$$\nabla f = 0 \iff f_{x_1} = f_{x_2} = f_{x_3} = 0$$

$$\Longrightarrow$$

$$f_{x_1} = 0 \iff 10x_1(x_2^2 + x_3^2) = \underbrace{2x_1(4x_2^2 + 3x_3^2)}_{\text{without } x_3 \text{ term}}$$

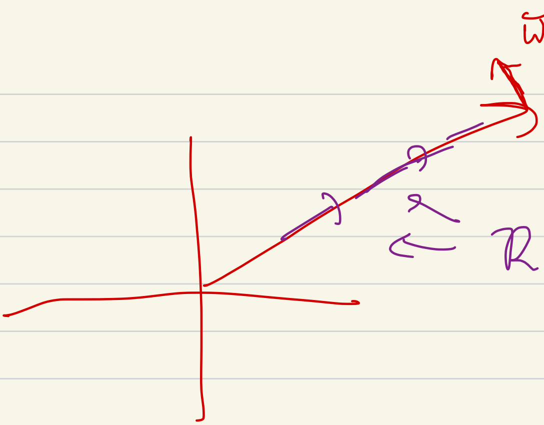
$$10x_1x_2^2 = 8x_1x_2^2$$

$\Rightarrow$ either x_1 or x_2 are 0

$$f_{x_2} = 0 \iff$$

$$f_{x_3} = 0 \iff$$

look at $\tilde{X} + \varepsilon \overline{w}$

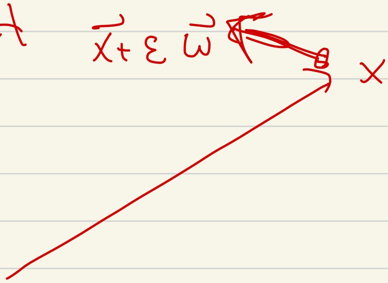


R_A is constant

$\vec{\omega} \perp \vec{x}$

Calculation gets simpler

but $\vec{x}+\varepsilon\vec{\omega}$



$\nabla f = 0$ at $\vec{x}$ $\iff$

for any $\vec{\omega}$,

$$\frac{d}{d\varepsilon} f(\vec{x}+\varepsilon\vec{\omega})\Big|_{\varepsilon=0} = 0$$

$$\begin{bmatrix} 5 & & \\ & 4 & \\ & & 3 \end{bmatrix}$$ largest eigenval.

because distinct

$f_{x_1} =$ either $x_1 = 0$

or $10(x_2^2 + x_3^2) = 8x_2^2 + 6x_3^2$

$$2x_2^2 + 4x_3^2 = 0$$

then

so $x_2 = x_3 = 0$

$$f(x) = f(x_0) + \left(\nabla f(x_0)\right)(x - x_0)$$

$$(x - x_0)^T \frac{1}{2} (\text{Hess} f)(x_0)(x - x_0)$$

+ higher order

$$\bar{d}(t) = \max_{\mu, \nu \text{ stoch}} \| \mu P^t - \nu P^t \|$$

nicest in that

$$\boxed{\bar{d}(t+s) \le \bar{d}(t)\,\bar{d}(s), \quad (*)}$$

(claim: max is attained at μ, ν = some standard basis vector)

In practice

$$d(t) = \max_{\mu} \| \mu P^t - \pi \|$$

stationary

We can say $(\pi P^t = \pi)$

$$\boxed{d(t) \leq \bar{d}(t) \underset{\text{triangle}}{\leq} 2\,d(t)}$$

$$\begin{aligned} \|\mu P^t - \nu P^t\| &\leq \|\mu P^t - \pi\| \\ &\quad + \|\pi - \nu P^t\| \end{aligned}$$

Mixing time definition uses $d(t)$

If $d(t_0) \leq \frac{1}{4} \Rightarrow \bar{d}(t_0) \leq \frac{1}{2}$, $\bar{d}(t_0 k) \leq \frac{1}{2^k}$

$d(t_0 k) \leq 1/2^k$.

$$\mu P^t - \nu P^t$$

$$= [\mu_1 \cdots \mu_n] P^t - [\nu_1 \cdots \nu_n] P^t$$

$$= [(\mu_1 - \nu_1)\ (\mu_2 - \nu_2) \cdots (\mu_n - \nu_n)] P^t$$

$$\| \quad \|_{TV} = \frac{1}{2} \| \quad \|_{L^1}$$

$$\max_{\mu,\nu} \| \mu P^t - \nu P^t \|_{L^1}$$

$$= \max_{\mu,\nu} \| (\mu_1 - \nu_1) \vec{e}_1 P^t + \cdots + (\mu_n - \nu_n) \vec{e}_n P^t \|_{L^1}$$

$$\|\mu P - \nu P\|_{TV} =$$

$$\max_{I \subset [n]} \sum_{i \in I} (\mu P - \nu P)_i$$

$$= \max_{I \subset [n]} \underbrace{\sum_{j \in I} \left((\mu_j - \nu_j) e_j^T P \right)_i}$$

$$= \sum_{i \in I} (\mu_i - \nu_i) \leq 1$$

$$\le \sum_{i \in I} \sum_{\sigma_j = \mu_j - \nu_j} \left(\sigma_j \, e_j^T P \right)_i$$

$$\le \left(\sum \sigma_j \right) \left(\sum_j \max_{j=1,\dots,n} \left(e_j^T P \right)_i \right)$$

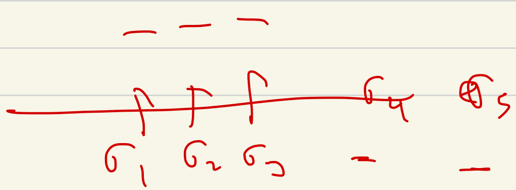
$$\max_{\mu,\nu} \underbrace{\|\mu P - \nu P\|_{L^1}}$$

$$\Big\|\sum_i (\mu_i - \nu_i)\, e_i^T P\Big\|_{L^1} \longrightarrow \Big\|\sum_i \sigma_i e_i^T P\Big\|_{L^1}$$

Let $\sigma_i = \mu_i - \nu_i$

$\sum \mu_i = 1, \quad \sum \nu_i = 1$

Let $\bar{I} = \{ i \mid \sigma_i \geq 0 \}$



$\sum \sigma_i = 0$

σ_i — — — $\swarrow$ $\sum_{i \in I} \sigma_i \leq 1$

$\swarrow$

— — — $\sum_{j \notin I} \sigma_j \geq -1$

[If $\mu = \nu$, $\sigma_i = 0$ for all i]

$=$

Say! $\sigma_i = \frac{1}{k}$, $i \in I$, $\sigma_j = \frac{-1}{k}$ $j \notin I$

match each σ_i with $-\sigma_j$

$$\left\| \sum_i \sigma_i e_i^T \rho \right\| = \left\| \sum_{i \in I} \sigma_i e_i^T \rho + \sum_{j \notin I} \sigma_j e_j^T \rho \right\|$$

$$\leq \quad k \cdot \frac{1}{k} \max_{\substack{i \in I \\ j \notin I}} \left\| (e_i^T - e_j^T) P \right\|_{1}$$

μ, ν

$$\mu_i - \nu_i \geq 0, \quad \text{say} \quad 0 \leq \sum_{i \in I} (\mu_i - \nu_i) \leq 1$$

$1 \quad \cdots \quad k$

$\mu_i - \nu_i < 0$

$$0 \geq \sum_{j \notin J} (\mu_j - \nu_j) \geq -1$$

$\{i \mid \sigma_i > 0\} = I$

$\{j \mid \sigma_j < 0\} = J$

$\boxed{\sigma_1 = \frac{1}{3}, \quad \sigma_2 = \frac{1}{3}, \quad \sigma_3 = \frac{1}{3}}$

$\boxed{\sigma_4 = \frac{-1}{2}, \quad \sigma_5 = -\frac{1}{2}}$

$$\| \mu P - \nu P \|_{L^1}$$

$$\left\| \frac{1}{3} e_1 P + \frac{1}{3} e_2 P + \frac{1}{3} e_3 P - \frac{1}{2} e_4 P - \frac{1}{2} e_5 P \right\|_{L^1}$$

$$\leq \max \| (e_i - e_j) P \| = M$$

$$\| \frac{1}{3} e_1 P - \frac{1}{3} e_4 P \| \leq \frac{1}{3} M$$

$$\left\| \frac{1}{3} e_2 P - \frac{1}{3} e_5 P \right\| \leq \frac{1}{3} m$$

$$\left\| \frac{1}{6} e_3 P - \frac{1}{6} e_4 P \right\| \leq \frac{1}{6} m$$

$$\left\| \frac{1}{6} e_3 P - \frac{1}{6} e_5 P \right\| \leq \frac{1}{6} m$$

$$\left\| \frac{1}{3} e_1 P + \frac{1}{3} e_2 P + \frac{1}{3} e_3 P - \frac{1}{2} e_4 P - \frac{1}{2} e_5 P \right\| \leq m$$

$$0 \leq \sum_{\substack{i \text{ s.t.} \\ \mu_i - \nu_i \geq 0}} (\mu_i - \nu_i) \leq 1$$

σ_i

$$\|$$

$$\overbrace{- \sum_{\substack{j \text{ s.t.} \\ \mu_j - \nu_j \leq 0}} (\mu_j - \nu_j)}$$

σ_i

$$\sigma_1 = \frac{1}{3} \quad \sigma_2 = \frac{1}{3}, \quad \sigma_3 = \frac{1}{10}$$

$$\sigma_4 = -\frac{1}{4} \qquad \sigma_5 = -\frac{1}{3} - \frac{1}{3} - \frac{1}{10} + \frac{1}{4}$$

$$\sigma_1, \sigma_2, \sigma_3 \geq 0, \qquad \sigma_1 + \cancel{\sigma_2} + \sigma_3 = \frac{2}{3} + \frac{1}{10} < 1$$

$$\sigma_4 + \sigma_5 = -\frac{2}{3} - \frac{1}{10}$$

If so

$$\|\mu P - \nu P\|_{TV} \leq \left(\max_{\substack{i=1,2,3 \\ j=4,5}} \|e_i P - e_j P\| \right) \cdot$$

L^1

$$\left(\frac{1}{3} + \frac{1}{3} + \frac{1}{10}\right)$$

$\mu = e_i$

any i, j.

$\nu = e_j$

$$\| \mu P - \nu P \|_{TV} = \| (e_i - e_j) P \|_{TV}$$

$\overset{\text{"}}{\leftarrow}$ $\overset{\text{"}}{\leftarrow}$

Also coupling argument

Connect, couple

$\sigma_1, \sigma_2, \sigma_3 > 0$ with

$\sigma_4, \sigma_5 < 0$