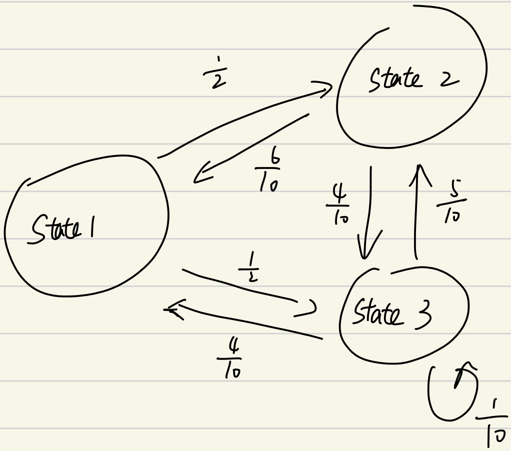
EECS 551 YUZHAN JIANG

Pr 1:

(a)



$$P = \begin{bmatrix} 0 & \frac{6}{10} & \frac{4}{10} \\ \frac{1}{2} & 0 & \frac{5}{10} \\ \frac{1}{2} & \frac{4}{10} & \frac{1}{10} \end{bmatrix}$$

This is Markov chain with $P \neq I$, π is an equilibrium distribution

(b) $\pi P = \pi$

$$\begin{cases} \pi_1 = \frac{6}{10}\pi_2 + \frac{4}{10}\pi_3 \\ \pi_2 = \frac{1}{2}\pi_1 + \frac{5}{10}\pi_3 \\ \pi_3 = \frac{1}{2}\pi_1 + \frac{4}{10}\pi_2 + \frac{1}{10}\pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

$\Rightarrow$ $\pi_1 = \frac{1}{3}$, $\pi_2 = \frac{1}{3}$ $\pi_3 = \frac{1}{3}$

π is the unique equilibrium distribution.

(c)

This P matrix is irreducible and aperiodic

$\Rightarrow$ P is primitive

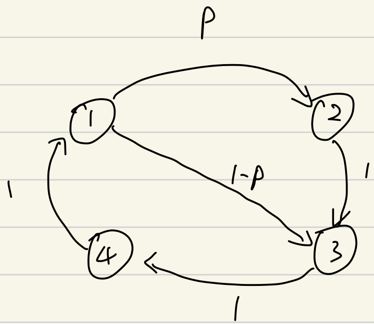
Therefore, the power iteration converges.

$\pi_{k+1} = P\pi_k$ is guaranteed to converge to π.

P2.

Let $N=4$.

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ p & 0 & 0 & 0 \\ 1-p & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$



$P\pi = \pi$

$$\Rightarrow \begin{cases} \pi_1 = \pi_4 \\ \pi_2 = p\pi_1 \\ \pi_3 = (1-p)\pi_1 + \pi_2 \\ \pi_4 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \end{cases}$$

$\pi_1 = \pi_3 = \pi_4$

$\Rightarrow \pi_1 + p\pi_1 + \pi_1 + \pi_1 = 1$

$\pi_1(3+p) = 1$

$\pi_1 = \frac{1}{3+p}$

$$\Rightarrow \pi_1 = \pi_3 = \pi_4 = \frac{1}{3+p}, \quad \pi_2 = \frac{p}{3+p}$$

So our conjecture: for $N > 4$

$$P = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ p & 0 & 0 & \cdots & \cdots & 0 \\ 1-p & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & 1 & 0 \end{bmatrix} \quad N \times N \qquad \pi = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \vdots \\ \pi_N \end{bmatrix}$$

$\Rightarrow$ ① $\pi_1 = \pi_3 = \pi_4 \cdots = \pi_N$

② $\pi_2 = p\pi_1$

③ $\pi_1 + \pi_2 + \cdots + \pi_N = 1$

$$\Rightarrow \pi_1 = \pi_3 = \cdots = \pi_N = \frac{1}{p+N-1}$$

$$\pi_2 = \frac{p}{p+N-1}$$

$$\therefore \pi_* = \begin{bmatrix} \frac{1}{p+N-1} \\ \frac{p}{p+N-1} \\ \vdots \\ \frac{1}{p+N-1} \end{bmatrix}_{N \times 1}$$

π_* is unique since P is non-negative matrix and strongly connected graph, so it is irreducible then the equilibrium distribution is unique.

P3:

convex set:

for any $x_1, x_2 \in C$ and $0 \le \theta \le 1$

if $\theta x_1 + (1-\theta) x_2 \in C$, then C is convex set

Let $P_1 = (x_1, r_1) \in B_r$ where $\|x_1\| \le r_1$

$P_2 = (x_2, r_2) \in B_r$ where $\|x_2\| \le r_2$

$P = \theta P_1 + (1-\theta) P_2 = (x, r)$

$\Rightarrow P = \theta (x_1, r_1) + (1-\theta)(x_2, r_2)$

$= (\theta x_1, \theta r_1) + ((1-\theta) x_2, (1-\theta) r_2)$

$= (\theta x_1 + (1-\theta) x_2, \theta r_1 + (1-\theta) r_2)$

$= (x, r)$

$\therefore$ $x = \theta x_1 + (1-\theta) x_2$ and $r = \theta r_1 + (1-\theta) r_2$

$$\begin{aligned} \|x\| = \|\theta x_1 + (1-\theta) x_2\| &\le \|\theta x_1\| + \|(1-\theta) x_2\| \\ &\le \theta \|x_1\| + (1-\theta) \|x_2\| \\ &\le \theta r_1 + (1-\theta) r_2 = r \end{aligned}$$

$\therefore$ $(x, r) \in B_r$

$\therefore$ B_r is convex set by the definition

P4:

(a) We rewrite $\varphi(t) \triangleq (\max(1-t,0))^2 \Rightarrow \varphi(t) = \begin{cases} (1-t)^2, & t \le 1 \\ 0, & t > 1 \end{cases}$

$$\therefore \nabla\varphi(t) = \begin{cases} -2(1-t), & t \le 1 \\ 0, & t > 1 \end{cases}$$

$$\begin{aligned} \Rightarrow \|\nabla\varphi(x) - \nabla\varphi(y)\|_2 &= \|2(x-1) - 2(y-1)\|_2 \\ &= \|2(x-y)\|_2 \\ &\le 2\|(x-y)\|_2 \quad \text{for} \quad x,y \le 1 \end{aligned}$$

Therefore, $L = 2$ for $x, y \le 1$

(b) Let $A = \begin{bmatrix} y_1 v_1^T \\ y_2 v_2^T \\ \vdots \\ y_m v_m^T \end{bmatrix}$

$$f(x) = 1_m' \cdot \varphi(Ax) + \beta \frac{1}{2}\|x\|_2^2$$

$$\begin{aligned} \nabla f(x) &= \sum_{m=1}^{M} \nabla\varphi(A_m : x) + \beta x \\ &= (\sum_{m=1}^{M} A_{m,:} \varphi(A_{m,:} x) + \beta x \\ &= A'\varphi(Ax) + \beta x \end{aligned}$$

$$\begin{aligned} \|\nabla f(x) - \nabla f(y)\|_2 &= \|A'\varphi(Ax) + \beta x - A'\varphi(Ay) - \beta y\|_2 \\ &= \|A'(\varphi(Ax) - \varphi(Ay)) + \beta(x-y)\|_2 \\ &\le \|A'(\varphi(Ax) - \varphi(Ay))\|_2 + \beta\|(x-y)\|_2 \\ &\le |||A'|||_2 \; L_\varphi \|Ax - Ay\|_2 + \beta\|(x-y)\|_2 \\ &\le L_\varphi |||A|||_2 \; |||A||| \; \|x-y\|_2 + \beta\|(x-y)\|_2 \\ &= (|||A'A|||_2 L_\varphi + \beta) \|x-y\|_2 \end{aligned}$$

$$\therefore L = 2|||A'A|||_2 + \beta$$

P3:

$$A = \begin{bmatrix} 20 & 14 & w & 17 \\ 44 & x & 16 & y \\ 8 & 6 & z & 7 \end{bmatrix}$$

$$A = i + \begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} d & e & f & g \end{bmatrix}$$

$$= i + \underbrace{\begin{bmatrix} ad & ae & af & ag \\ bd & be & bf & bg \\ cd & ce & cf & cg \end{bmatrix}}$$ (Let call this B)

$$= \begin{bmatrix} i+ad & i+ae & i+af & i+ag \\ i+bd & i+be & i+bf & i+bg \\ i+cd & i+ce & i+cf & i+cg \end{bmatrix}$$

① $i+ad - (i+ag) = a(d-g) = 3$

② $i+cd - (i+cg) = c(d-g) = 1$

$\frac{①}{②} = 3 = \frac{a}{c}$

$a = 3c$

And we know

$i+ag = 17 \quad i+cg = 7$

$ag - cg = 10 = 2cg = 10$

$cg = 5 \quad \therefore i = 2$

$$B = \begin{bmatrix} 18 & 12 & w-2 & 15 \\ 42 & x-2 & 14 & y-2 \\ 6 & 4 & z-2 & 5 \end{bmatrix}$$

$ae = 12$

$be = x-2 = 4 \times 7 = 28 \quad x = 2$

$ce = 4$

$z-2 = \frac{14}{7} = 2$

$z = 4$

$y-2 = 5 \times 6$

$y = 32$

$w-2 = 2 \times 3 = 6$

$w = 8$

Overall, $x = 2$, $y = 32$, $z = 4$, $w = 8$

P6:

(a)

The eigenvalue of $I - \alpha A'A$

$$\mathrm{eig}(I - \alpha A'A) = 1 - \alpha\, \mathrm{eig}(A'A)$$
$$= 1 - \alpha\, \sigma_1^2(A) \le 1 - \alpha\, \sigma_2^2(A) \cdots \le 1 - \alpha\, \sigma_N^2$$

$\therefore$ $\rho(I - \alpha A'A) = \max\left(|1 - \alpha \sigma_1^2(A)|, \ |1 - \alpha \sigma_N^2(A)| \right)$

$\therefore$ If we want to minizing $\rho(I - \alpha A'A)$:

the step size $\alpha_* = \dfrac{2}{\sigma_1^2(A) + \sigma_N^2(A)}$

(b) $\rho(I - \alpha_* A'A) = 1 - \alpha_* \sigma_N^2(A)$

$$= 1 - \frac{2}{\sigma_1^2(A) + \sigma_N^2(A)} \sigma_N^2$$

$$= \frac{\sigma_1^2(A) - \sigma_N^2(A)}{\sigma_1^2(A) + \sigma_N^2(A)}$$

PT:

(a) $x_{k+1} = x_k - \alpha P_0 A'(Ax - y)$, this is governed by the eigenvalue of $I - \alpha P^{\frac{1}{2}} A'A P^{\frac{1}{2}}$

$$\alpha_* = \frac{2}{\sigma_1^2(AP_0^{\frac{1}{2}}) + \sigma_N^2(AP_0^{\frac{1}{2}})}$$

$$= \frac{2}{\sigma_1(P_0^{\frac{1}{2}}A'AP^{\frac{1}{2}}) + \sigma_N(P_0^{\frac{1}{2}}A'AP_0^{\frac{1}{2}})}$$

(b) $P_0 = (A'A)^{-1}$

Thus, $P_0^{\frac{1}{2}} A'A P_0^{\frac{1}{2}}$

$= P_0^{\frac{1}{2}} P_0^{-1} P_0^{\frac{1}{2}}$

$= I$

$\therefore \alpha_* = \frac{2}{\sigma_1 + \sigma_N}$