Advanced Linear Algebra Week 5 Day 1

2018/10/08 – Jonathan Hayase, updated by Prof. Weiqing Gu

1 Positive Matrices and Applications to Markov Processes

Definition: A matrix $P = \left(p_{ij}\right)_{n\times n}$ is positive if if $p_{ij} \in \mathbb{R}$ and $p_{ij} > 0$ for all $i, j \in 1, \ldots, n$.

Note: A matrix being positive is not the same as being positive definite (which is $x^T P x > 0$).

Notation:

1. For $x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T$ and $y = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^T$, we say $x < y$ if $x_i < y_i$ for all $i = 1, \ldots, n$.
2. Say $x < y$ if $x_i \le y_i$ as above.

   Note that tripotomy does not hold here, so $x \le y \;\not\Rightarrow\; x < y \vee x = y$.
3. Let $\xi_0 = (1, 1, \ldots, 1) \in R$ and vector $x > 0$. We say x is L_1-normalized if

$$\xi_0 = (1, \ldots, n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i = 1.$$

1.1 Perron's Theorem

If P is a positive matrix, then P has a dominant eigenvalue $\lambda(P)$ such that:

1. $\lambda(P) > 0$, $\lambda(P)$ is an eigenvalue of P and there exists a eigenvector $h > 0$ such that $Ph = \lambda(P)h$.
2. $\lambda(P)$ is a simple (e.g. multiplicity 1) eigenvalue.
3. For any eigenvalue μ, then $|\mu| \le \lambda(P)$.
4. For any eigenvalue μ, then if $\mu \ne \lambda(P)$ we have $|\mu| < \lambda(P)$ and (u, f) is an eigen pair, then $f \not\geq 0$.

1.1.1 Proof

Idea: Want to maximize λ such that $Px = \lambda x$ and $x \ne 0$ and $x \ge 0$ for positive P. Define

$$\mathscr{N}(P) = \left\{\lambda \in R \mid \lambda \ge 0,\ \exists x \subseteq \mathbb{R}^n,\ x \ge 0,\ \xi_0 x = 1,\ P\lambda \ge \lambda x\right\} \subseteq \mathbb{R}$$

Goal is to show that $\mathscr{N}(P)$ is nonempty, bounded, and closed subset of $\mathbb{R}$.

First, we show that $\mathscr{N}(P)$ is non-empty. Clearly, $0 \in \mathscr{N}(P)$, since we may take $x = (1, 0 \ldots)$ so $\xi_0 x = 1$ and observe that $Px \ge 0x \ge 0$. Moreover, $\mathscr{N}(P)$ contains nonzero element. Take $y =$ any positive vector, then $Py > 0$. Let $x = {}^{y}/_{\xi_0 y} > 0$, then $\xi_0 x = \xi_0 \left({}^{y}/_{\xi_0 y}\right) = 1$ and $Px > 0 = \lim_{x\to 0} \lambda x$. Notice that $\lambda \ge 0$ so $\lim_{\lambda\to 0} \lambda x = 0$ so there must exist $\lambda 0 > 0$ such that $\lambda_0 x \le Px$. Thus, $\lambda_0 \in \mathscr{N}(P)$.

Second, we show that $\mathscr{N}(P)$ is bounded. Let $\lambda \in \mathscr{N}(P)$. Then there exists $x \in \mathbb{R}^n$ such that $x \geq 0$, $\xi_0 x = 1$, and $Px \geq \lambda$. Thus, $\xi_0 Px \geq \xi_0 \lambda x = \lambda \xi_0 x = \lambda$. Now write $\xi_0 P = (b_1, \ldots, b_n)$ where b_i is the sum of the i_{th} column of P. Let $b = \max_i (b_1, \ldots, b_n)$. Then $b\xi_0 \geq \xi P$, so

$$b = b1 = b\left(\xi_0 x\right) = \left(b\xi_0\right) x \geq \xi_0 Px \geq \lambda$$

Since $\lambda \geq 0$, we have $\lambda \in [0, b]$ for any $\lambda \in \mathscr{N}(P)$. Hence, $\mathscr{N}(P)$ is a bounded subset of $\mathbb{R}$.

Aside (Stochastic Matrices): Suppose you have three lily pads, and the probability that the frog jumps from pad 1 to pad 2 is $^2/_3$ and the probability that the frog stays or jumps to pad 3 are both $^1/_6$. We can construct a matrix that describes this sort of behavior:

$$P = \begin{bmatrix} ^1/_6 & ^2/_3 & ^1/_6 \\ ^1/_2 & 0 & ^1/_2 \\ ^1/_4 & ^1/_2 & ^1/_4 \end{bmatrix}$$

is a stochastic matrix, P is positive, and the sum of the rows of P equal 1.

Note: If P is a stochastic matrix, then $\xi_0 P = (1, 1, \ldots, 1)$ so $b = 1$, so $0 \leq \lambda \leq 1$.

Third, we are going to show that $\mathscr{N}(P)$ is a closed subset of $\mathbb{R}$. Let $\{\lambda_m\}_{m=1}^{\infty}$ be a sequence in $\mathscr{N}(P)$ and $\lim_{m\to\infty} \lambda_m = \lambda$. We wish to show that $\lambda \in \mathscr{N}(P)$. Details of this will be handed out.

Claim is proved, so there exists $\lambda(P) = \max\left\{\lambda \subseteq \mathscr{N}(P)\right\} > 0$. Moreover, there exists an $h \in \mathbb{R}^n$ such that $h \geq 0$, $\xi_0 h = 1$, and $Ph \geq \lambda(P)h$.

Claim 2: $Ph = \lambda(P)h$ so $\lambda(P)$ is an eigenvalue w/ non-negative eigenvectors. Idea of proof: Prove by contradiction. Details in handout.

Claim 3: In fact, we can show that $h > 0$. Proof: Since $P > 0$ and $h \geq 0$ and $h \neq 0$, then $Ph > 0$. Now, since $Ph = \lambda(P)h$ implies $\lambda(P)h > 0$. So $h > 0$ since $\lambda(P) > 0$.

Claim 4: $\lambda(P)$ is a simple eigenvalue. Detail will be handed out.

Claim 5: For any eigenvalue k of P, then $|k| \leq \lambda(P)$ and $|k| = \lambda(P)$ iff $k = \lambda(P)$. Proof: Let (k, y) be an eigen pair, $y = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^T \neq 0$. Then $Py = ky$. Look at the s^{th} element of above equation.

$$\sum_{i=1}^{n} P_{sj} y_j = k y_s$$

$$|k||y_s|j = |ky_s| = \left|\sum_{j=1}^{n} p_{sy} y_j\right| \leq \sum_{j=1}^{n} \left|p_{sj} y_j\right| = \sum_{j=1}^{n} P_{sj} \left|y_j\right|$$

Let $z = \left(|y_1|, \ldots, |y_n|\right) \geq 0$. Then we have $Pz \geq |k|\, z$. Let $c = {}^{z}/_{\xi_0 z}$, then $x \geq 0$, $\xi_0 x = 1$, and $Px \geq |k|\, x$. Therefore, $|k| \in \mathscr{N}(P)$, so $|K| \leq \lambda(P)$. Next, we can show if $|k| = \lambda(P)$ then $k = \lambda(P)$. Details in handout.

Claim 6: Suppose that $k < \lambda(P)$ and (k, f) is an eigen pair, then $f \ngeq 0$. Proof: consider P^T we also have $P^T > 0$. Then, (Exercise) $\lambda(P) = \lambda(P^T)$. Let $\xi > 0$ such that $(\lambda(P), \xi)$ is an eigen pair for P^T. Now, $Pf = kf$ and $f \neq 0$. $P^T \xi = \lambda(P)\xi$, so $(P^T \xi)^T = \lambda(P)\xi^T$. Acting on f implies $\xi^T P f \cdot \lambda(P)\xi^T f$. Multiplying both sides by ξ^T yields $\xi^T pf = k\xi^T$ so $\left(\lambda(P - k)\right) \xi^T f = 0$ because $\lambda(P) - k \neq 0$, so $\xi^T f = 0$ and $\xi > 0$ so $\xi^T f = 0$ implies $f \ngeq 0$.

2 PageRank

$$\begin{bmatrix} 1/4 & 0 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ 1/4 & 1/2 & 0 \end{bmatrix} \to \begin{bmatrix} 1/4+p & p & 1/2+p \\ 1/2+p & 1/2+p & 1/2+p \\ 1/4+p & 1/2+p & p \end{bmatrix}$$

Can use Perron's Theorem.

As we noticed that for a Markov chain, the transition stochastic matrix P may not be positive. But $P \geq 0$. (P is regular if there exist an $m \in \mathbb{N}$ such that $P^m > 0$). Consider the example $P = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, then P is not regular, since $P^2 = I$, but $P^2 = P$, and this pattern repeats. We still would like to analyze this situation.

2.1 Frobenius Theorem

Let $P \geq 0$ be a $n \times n$ matrix. Then there exist $\lambda(P)$ such that

1. $\lambda(P)$ is an eigenvalue of P, $\lambda(P) \geq 0$ and there exists $h \geq 0$ such that $\xi_0 h = 1$ and $Ph = \lambda(P)h$.
2. If $k \in \mathbb{C}$ is an eigenvalue of P, then
   (i) $|k| \leq \lambda(P)$.
   (ii) $|k| = \lambda(P)$ implies $k = \exp\left(\frac{2\pi ik}{m}\right)\lambda(P)$ for some $k, m \in \mathbb{N}$ and $m \leq n$.

2.1.1 Proof of 1

Let's perturb P. Let

$$P_m = \begin{bmatrix} 1/m & \cdots & 1/m \\ \vdots & \ddots & \vdots \\ 1/m & \cdots & 1/m \end{bmatrix} + P > 0$$

Then $\lim_{m\to\infty} P_m = P$. By Perron's Theorem, P_m has a dominant eigenvalue $\lambda(P) > 0$. By Theorem 6, P101, $\lim_{m\to\infty} \lambda(P_m)$ exists and is an eigenvalue of P.

Claim: $\lambda(P) \geq 0$ and there exists $h \geq 0$ such that $\xi_0 h = 1$ and $Ph = \lambda(P)h$. Because $\lambda(P_m) > 0$ we have $\lim_{m\to\infty} \lambda(P_m) \geq 0$ so $\lambda(P) \geq 0$.

Let h_m satisfy $P_m h_m = \lambda(P_m)h_m$, $h_m > 0$, and $\xi_0 h_m = 1$.

As we showed earlier, $\|h_m\| \leq \sqrt{n}$. Therefore, there exists a subsequence $\{h_{m_k}\}$ of $\{h_m\}$ such that $\lim_{k\to\infty} h_{m_k} \to h$. Because $h_{m_k} > 0$, we have $h \geq 0$.

Moweover since $\xi_0 h_{m_k} = 1$ taking the limit of both sides yields $\xi_0 h = 1$ so $h \neq 0$.

Now

$$\begin{tikzcd} P_{m_k} h_{m_k} \arrow[r, "="] \arrow[d, "m\to\infty"] & \lambda(P_{m_k}) h_{m_k} \arrow[d, "m\to\infty"] \\ Ph \arrow[r, "="] & \lambda(P)h \end{tikzcd}$$

2.1.2 Proof of 2

(i) Let $k \in \mathbb{C}$ be any eigenvalue of P. Then again by Theorem 6, P101, there exists an eigenvalue k_m of P_m such that $\lim_{m\to\infty} k_m = k$.

By Perron's Theorem,

$$\begin{tikzcd}
|k_m| \arrow[r, "\leq"] \arrow[d, "m\to\infty"] & \lambda(P_m) \arrow[d] \\
|k| \arrow[r, "\leq"] & \lambda(P)
\end{tikzcd}$$

(ii) $|k| = \lambda(P)$ so $|k| = e^{i\theta}\lambda(P)$. Then we show that $\theta = \exp\left(\frac{2\pi ik}{m}\right)$. Too much detail. Not required for this class.

3 Riesz Representation Theorem

Recall the Riesz Representation Theorem.

If V^n is an n dimensional vector space over a field K, either $\mathbb{R}^n$ or $\mathbb{C}^n$. Then V^* is the set of linear functions on V.

$$V^* = \big\{\ell \mid \ell : V \to K, \ell \text{ is linear}\big\}$$

For all $\lambda \in V^*$, there exists a y such that

$$\ell(x) = \langle x, y\rangle$$

for all x. If and only if $V^{**} = V$.

Key idea: if $\big(V, \langle\cdot,\cdot\rangle\big)$ then orthonormal because great/fast in use

$$v = \sum_{i=1}^{n} a_i u_i$$

But if we have no inner product, we want to mimic an orthonormal basis. We define a dual basis $\ell_i(e_j) = \delta_{ij}$. Mimics $u_i(uj) = \langle u_i, u_j\rangle = \delta_{ij}$.

If there exists an inner product on V^* then $\{\ell_1, \dots, \ell_n\}$ is $\{u_1, \dots, u_n\}$.

Now, we want to extend this to $C[a, b]$. But to do this we will need the following theorem:

3.1 Hahn–Banach theorem

Recall $\big(V, \langle\cdot,\cdot\rangle\big)$. Only having $\big(V, \|\cdot\|\big)$ is weaker. Even weaker is (V, p), where p is sublinear.

Definition: A sublinear functional is a real valued function p on vector space X which satisfies

1. $p(x + y) \leq p(x) + p(y)$ for $x, y \in X$.
2. $p(ax) = ap(x)$ for $a \geq 0, x \in X$.

3.1.1 Examples

For any linear map $\ell : X \to \mathbb{R}$, define $p(x) = \big|\ell(x)\big|$.

1. $p(x+y) = \big|\ell(x+y)\big| \leq \big|\ell(x)\big| + \big|\ell(y)\big| = p(x) + p(y)$.

2. $p(ax) = \big|\ell(ax)\big| = a\big|\ell(x)\big| = ap(x)$.

3.2 Statement of Hahn–Banach theorem for real vector spaces

Suppose we have $\big(X/\mathbb{R}, p\big)$ for sublinear p

$$f\Bigg|_{Z}^{\text{linear functional}} \implies \exists\, \tilde{f}\Bigg|_{X}^{\text{linear functional}} : \begin{cases} \tilde{f}(z) = f(z) & \text{for } z \in Z \\ \tilde{f}(x) \leq p(x) & \text{for all } x \in X \end{cases}$$

where $Z \subseteq X$ is a subspace of X if $f(x) \leq p(x)$ on Z.

Let x be a real vector space and p be a sublinear functional on X. Let f be a linear functional defined on a subspace $Z \subseteq X$ satisfying $f(z) \leq p(x)$ for all $z \in Z$.[1] Then, there exists linear functional $\tilde{f}$ satisfying

1. $\tilde{f}(z) = f(z)$ for all $z \in Z$, and

2. $\tilde{f}(x) \leq p(x)$ for all $x \in X$.

3.3 Proof

By Zorn's Lemma.

Recall: **Zorn's Lemma**: Let P be a partially ordered set such that every chain has an upper bound in P. Then the set P contains at least one maximum element.

Step 1: Form a set {all linear functionals of f that are dominated by p}.

$$S = \big\{\phi \mid \phi : \mathrm{D}_\phi \to \mathbb{R} \text{ is linear}, \mathrm{D}_\phi \supset Z, \phi = f \text{ on } z, \phi \leq p \text{ on } \mathrm{D}_\phi\big\}$$

Step 2: Define binary relation $<$ (extension) Say $\phi_1 < \phi_2$ if ϕ_2 is an extension of ϕ_1.

Step 3: Every chain $C \subseteq S$ has an upper bound.

[1] This line seems a bit suspicious.