Notes for Log-Concave Polynomials II

Xiaoyu Chen

1 背景知识

这篇文章中涉及了大量我以往不太了解的技术, 这里集中做一个整理. 为了节约时间, 我会略过所有我能马上反应过来的推导. 并且, 为了简化工作, 尽量省略了所有的交叉引用, 如果发现后面有地方的证明看不过去, 可以去翻前面的结论.

1.1 线性代数

1.1.1 Schur Product Theorem

Fact 1. $\mathrm{tr}(AB^T) = \sum_{i,j=1}^{n} A_{ij}B_{ij}$

proof. $\mathrm{tr}(AB^T) = \sum_i (AB^T)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}B^T_{ji} = \sum_{i,j=1}^{n} A_{ij}B_{ij}$ □

Fact 2. $x^*(A \circ B)y = \mathrm{tr}((\mathrm{diag}\ \overline{x})A(\mathrm{diag}\ y)B^T)$

proof. 首先可以发现: $[(\mathrm{diag}\ \overline{x})A]_{ij} = \sum_{k=1}^{n}(\mathrm{diag}\ \overline{x})_{ik}A_{kj} = \overline{x}_i A_{ij}$. 同理有: $[B(\mathrm{diag}\ y)]_{ij} = B_{ij}y_j$. 所以:

$$\begin{aligned}
\mathrm{tr}((\mathrm{diag}\ \overline{x})A(\mathrm{diag}\ y)B^T) &= \sum_{i,j=1}^{n} [(\mathrm{diag}\ \overline{x})A]_{ij}[B(\mathrm{diag}\ y)]_{ij} \\
&= \sum_{i,j=1}^{n} \overline{x}_i A_{ij}B_{ij}y_j \\
&= \sum_{i,j=1}^{n} \overline{x}_i (A \circ B)_{ij} y_j \\
&= x^*(A \circ B)y
\end{aligned}$$ □

Fact 3. *如果矩阵 B 是一个 Hermitian 阵, 则 $B^{1/2}$ 存在.*

proof. 因为 B 是一个 Hermitian 阵, 所以存在 P, 使得 $B = P\Lambda P^{-1}$(Λ 是一个对角阵). 这个时候显然, $B^{1/2} = P\Lambda^{1/2}P^{-1}$. □

Fact 4. $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$

proof.

$$\mathrm{tr}(ABC) = \sum_i [ABC]_{ii} = \sum_{i,j,k} A_{ij}B_{jk}C_{ki} = \sum_{i,j,k} C_{kj}A_{ij}B_{jk} = \mathrm{tr}(CAB)$$

□

Theorem 1 (Schur Product Theorem)**.** *如果 $A, B \in \mathbb{R}^{n\times n}$ 都是半正定的，则 $A \circ B$ 也是半正定的.*

△ 注: $A \circ B_{ij} = A_{ij}B_{ij}$

△ 注: 当人们在讨论正定或者半正定的矩阵的时候，一般都默认这些矩阵是 *Hermitian* 阵 (这个约定可以在 *wiki* 上关于正定矩阵的定义发现)

proof.

$$
\begin{aligned}
x^*(A \circ B)x &= \mathrm{tr}((\mathrm{diag}\ \overline{x})A(\mathrm{diag}\ x)B^T) \\
&= \mathrm{tr}((\mathrm{diag}\ \overline{x})A(\mathrm{diag}\ x)\overline{B}), \quad B \text{ is hermitian} \\
&= \mathrm{tr}(\overline{B}^{1/2}(\mathrm{diag}\ \overline{x})A(\mathrm{diag}\ x)\overline{B}^{1/2}) \\
&= \mathrm{tr}(C^*AC), \quad 其中 C = (\mathrm{diag}\ x)\overline{B}^{1/2}
\end{aligned}
$$

因为 A 半正定，所以令 $y = Cx$ 可得 $y^*Ay \geq 0$. 这也说明 C^*AC 半正定，所以 $\mathrm{tr}(C^*AC) \geq 0$. 所以 $x^*(A \circ B)x \geq 0$, 所以 $A \circ B$ 半正定. □

1.1.2 Positive Matrix

下面来学习一些关于 positive matrix 的结果.

Define 1 (positive matrix)**.** *如果对于一个矩阵 A, 有 $A_{ij} > 0, \forall i, j$, 则说 A 是一个 positive matrix.*

△ 注: 如果 A 是正项矩阵，计作 $A > 0$. 如果 x 是正项向量，计作 $x > 0$.

Define 2 (谱半径)**.** *矩阵 A 的谱半径定义为 $\rho(A) = \max\{|\lambda| : \lambda$ 是 A 的一个特征值$\}$*

Define 3. *令 x 是一个向量，则 $|x|$ 是一个向量，且 $|x|_i = |x_i|$.*

△ 注: 这里在复数域 $\mathbb{C}$ 上，所以对于 $a \in \mathbb{C}$, $|a|$ 指的是 a 的模长.

Fact 5. *对于 $\forall x \in \mathbb{C}^n$, 都 $\exists \theta \in \mathbb{R}$, 使得 $e^{-i\theta}x = |x|$.*

证明. 这个结论是复数乘法的几何意义的一个应用. 欧拉公式告诉我们，两个复数相乘，结果是模长相乘，夹角相加. 因为所有的复数都能被写成 $re^{i\theta}$ 的形式，其中 r 是模长，θ 是夹角. 所以，如果我们想知道一个复数 x 的模长，我们可以通过将其旋转到实轴正半轴来实现，旋转的操作可以有某个单位长度的复数来做到，具体来说，如果 $x = re^{i\theta}$, 我们只需要给他乘上一个 $e^{-i\theta}$ 即可. □

Fact 6. *$x \in \mathbb{C}^n$, $|\sum\limits_i x_i| \leq \sum\limits_i |x_i|$. 当 $\exists \theta \in \mathbb{R}$, 使得 $e^{-i\theta}x = |x|$ 时取等.*

△ 注: 这个式子可以看成绝对值不等式在复数域和高维向量空间的推广.

证明. 首先, 不难发现, $\sum_i x_i \in \mathbb{C}$. 所以由上面的结论, 就 $\exists\theta$, 使得 $e^{-i\theta}(\sum_i x_i) = |\sum_i x_i|$. 所以

$$
\begin{aligned}
|\sum_i x_i| &= e^{-i\theta}(\sum_i x_i) \\
&= \mathrm{Re}\left[e^{-i\theta}(\sum_i x_i)\right] \\
&= \sum_j \mathrm{Re}\left[e^{-i\theta}x_j\right] \\
&\leq \sum_j |e^{-i\theta}x_j| \\
&= \sum_j |x_j|
\end{aligned}
$$

从上面的式子, 不难发现, 这个式子如果要取到等号, 则 $\mathrm{Re}\left[e^{-i\theta}x_j\right] = |x_j|, \forall j$. 这个时候显然就有 $e^{-i\theta}x_j = |x_j|, \forall j$. 也就是说, 如果要取等号, 则有 $e^{-i\theta}x = |x|$. □

Fact 7. $|Ax| \leq |A|\,|x|$

证明.

$$
\begin{aligned}
|Ax|_k &= |\sum_i A_{ki}x_i| \\
&\leq \sum_i |A_{ki}x_i| \\
&= \sum_i |A_{ki}|\,|x_i| = [|A|\,|x|]_k
\end{aligned}
$$

□

Lemma 1. 令 A 是一个 *positive matrix*. λ, x 是 A 的一个特征值对, 且 $|\lambda| = \rho(A)$. 则有 $|x| > 0$ 且有 $A|x| = \rho(A)|x|$, 且 $\exists\theta$ 使得 $e^{i\theta}x = |x|$.

proof. 因为 $A > 0$, 所以直接就有 $z = A|x| > 0$. 所以

$$
z = A|x| = |A||x| \geq |Ax| = |\lambda x| = |\lambda||x| = \rho(A)|x|
$$

. 即 $z = A|x| \geq \rho(A)|x|$. 这个时候不妨设 $y = z - \rho(A)|x| \geq 0$.

(1) 如果 $y = 0$, 则 $\rho(A)|x| = A|x| > 0$, 所以 $\rho(A) > 0, |x| > 0$.

(2) 如果 $y > 0$, 则 $Ay = Az - \rho(A)z > 0$. 所以 $Az > \rho(A)z$. 这和 $\rho(A)$ 的定义矛盾了.

△ 注: $\lambda_{\max} = \max\limits_{x\neq 0} \dfrac{x^*Ax}{x^*x}$, 所以上面 (2) 中的现象会导致 $\dfrac{z^*Az}{z^*z} > \rho(A)$, 即 $\lambda_{\max} > \rho(A)$.

综合 (1), (2) 两点, 有 $A|x| = \rho(A)|x|$, 且 $\rho(A) > 0, |x| > 0$. 从 $A|x| = \rho(A)|x|$, 可以知道

$$
|\sum_i A_{ki}x_i| = \sum_i |A_{ki}x_i|, \quad \forall k
$$

这时对于某个固定的 k, 有 $\exists\theta \in \mathbb{R}$, 使得

$$
e^{-i\theta}A_{kj}x_j = |A_{kj}||x_j|, \quad \forall j
$$

因为 $A > 0$, 所以

$$e^{-i\theta} x_j = |x_j|, \quad \forall j$$

, 即 $e^{-i\theta} x = |x|$. □

Lemma 2. *$A \in \mathbb{R}^{n\times n}, A > 0$, 则有, A 关于 $\rho(A)$ 的几何重数是 1.*

proof. 设有 $w, z \in \mathbb{C}^n$, 且有 $\substack{Aw=\rho(A)w \\ Az=\rho(A)z}$. 由前面的结论可知:

(1) $\exists \theta_1 \in \mathbb{R}$ 使得 $p = e^{-i\theta_1} z > 0$, 且 $Ap = \rho(A)p$.

(2) $\exists \theta_2 \in \mathbb{R}$ 使得 $q = e^{-i\theta_2} w > 0$, 且 $Aq = \rho(A)q$.

这个时候, 不妨设 $\beta = \min\limits_{1\le i\le n} \dfrac{q_i}{p_i}$, $r = q - \beta p$. 这时 $r \ge 0$ 且 r 至少有一个地方为 0.

如果 $r \ne 0$, 则

$$0 < Ar = A(q - \beta p) = \rho(A) q - \beta \rho(A) p = \rho(A) r$$

所以 $\rho(A) > 0, r > 0$, 这和 r 上有一个地方为 0 有矛盾.

若果 $r = 0$, 则 $q = \beta p$.

$$\begin{aligned} e^{-i\theta_2} w &= \beta e^{-i\theta_1} z \\ w &= \beta e^{-i(\theta_1 - \theta_2)} z \end{aligned}$$

故 w 和 z 只呈常数倍关系. □

将上面两个引理的结论收集起来, 就能得到下面定理.

Theorem 2 (Perron-Frobenius Theorem)**.** *令 $A \in \mathbb{R}^{n\times n}, A > 0$. 则 A 有一个严格为正的特征值 $\lambda = \rho(A)$. 且 λ 对应的特征向量 $x > 0$. A 关于 λ 的几何重数是 1.*

1.1.3 特征值和子空间中的向量

Fact 8. *对于一个矩阵 A, 我们可以将他的特征值从小打到排列:*

$$\lambda_{\min} = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_{n-1} \le \lambda_n = \lambda_{\max}$$

Fact 9. *对于 Hermitian 矩阵来说, 我们有 n 个相互正交的特征向量*

Theorem 3 (Rayleigh theorem)**.** *假设有 $1 \le i_1 \le \cdots \le i_k \le n$. 且有 $x_{i_1}, x_{i_2}, \cdots, x_{i_k} \in \mathbb{C}^n$ 相互正交. 且有 $Ax_{i_p} = \lambda_{i_p} x_{i_p}, p = 1, 2, \cdots, k$. 然后令 $S = \mathrm{Span}\{x_{i_1}, x_{i_2}, \cdots x_{i_k}\}$ 则有:*

$$\begin{aligned} \lambda_{i_1} &= \min_{\substack{x\ne 0\\ x\in S}} \frac{x^*Ax}{x^*x} = \min_{\substack{x\in S\\ ||x||_2=1}} x^*Ax \\ &\le \max_{\substack{x\in S\\ ||x||_2=1}} x^*Ax = \max_{\substack{x\ne 0\\ x\in S}} \frac{x^*Ax}{x^*x} = \lambda_{i_k} \end{aligned}$$

证明. 对于 $\forall x \in S, ||x||_2 = 1$, 我们有:

$$x = \alpha_1 x_{i_1} + \alpha_2 x_{i_2} + \cdots + \alpha_k x_{i_k}$$

且

$$\begin{aligned} x^*x &= 1 \\ 1 &= \sum_{j=1}^{k} (\alpha_j x_{i_j})^* \alpha_j x_{i_j} \\ &= \sum_{j=1}^{k} \overline{\alpha_j} \alpha_j x_{i_j}^* x_{i_j} \\ &= |\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_k|^2, \quad \text{这里不妨 } x_{i_j} \text{ 都是单位长度向量} \end{aligned}$$

又有

$$\begin{aligned} x^*Ax &= (\alpha_1 x_{i_1} + \cdots + \alpha_k x_{i_k})^* (\alpha_1 \lambda_{i_1} x_{i_1} + \cdots \alpha_k \lambda_{i_k} x_{i_k}) \\ &= |\alpha_1|^2 \lambda_{i_1} + \cdots + |\alpha_k|^2 \lambda_{i_k} \end{aligned}$$

因为 $|\alpha_1|^2 + \cdots + |\alpha_k|^2 = 1$, 所以 x^*Ax 是 $\lambda_{i_1}, \cdots, \lambda_{i_k}$ 的一个凸组合. 所以, 当然有 $\lambda_{i_1} \leq x^*Ax \leq \lambda_{i_k}$, 并且刚好可以取到端点. □

Lemma 3 (空间交)**.** 令 $S_1, S_2, \cdots, S_k$ 是 $V = \mathbb{C}^n$ 的子空间.
如 $\delta = \dim S_1 + \dim S_2 + \cdots \dim S_k - (k-1)n \geq 1$, 则 $S_1 \cap S_2 \cap \cdots \cap S_k$ 中有 δ 个相互正交的单位向量.

证明. 先来考虑两个子空间的情形. 对于任意两个子空间 $S_1, S_n \subset V$, 有

$$\dim (S_1 \cap S_2) + \dim (S_1 + S_2) = \dim S_1 + \dim S_2$$

所以

$$\begin{aligned} \dim (S_1 \cap S_2) &= \dim S_1 + \dim S_2 - \dim (S_1 + S_2) \\ &\geq \dim S_1 + \dim S_2 - \dim V \\ &= \dim S_1 + \dim S_2 - n \end{aligned}$$

然后, 在这个基础上, 我们来使用归纳法. 假设对于 k 个子空间, 有:

$$\dim (S_1 \cap \cdots \cap S_k) = \sum_{i=1}^{k} \dim S_i - (k-1)n$$

则对于 $k+1$ 个子空间的情况, 我们有:

$$\dim (S_{k+1} \cap (S_1 \cap \cdots S_k)) + \dim (S_{k+1} + (S_1 \cap \cdots \cap S_k)) = \dim (S_1 \cap \cdots \cap S_k) + \dim S_{k+1}$$

所以

$$
\begin{aligned}
\dim(S_1\cap\cdots\cap S_{k+1}) &= \sum_{i=1}^{k}\dim S_i-(k-1)n+\dim S_{k+1}-\dim(S_{k+1}+(S_1\cap\cdots\cap S_k))\\
&\le \sum_{i=1}^{k+1}\dim S_i-(k-1)n-\dim V\\
&= \sum_{i=1}^{k+1}\dim S_i-(k-1)n-n\\
&= \sum_{i=1}^{k+1}\dim S_i-kn
\end{aligned}
$$

所以这里有 $\dim(S_1\cap\cdots S_k)\ge\delta$. □

Theorem 4 (Courant-Fischer Theorem)**.** *A 是 Hermitian 阵, 则*

$$\lambda_k=\min_{\substack{S\subset V\\ \dim S=k}}\max_{\substack{x\neq 0\\ x\in S}}\frac{x^*Ax}{x^*x}$$

proof. 由题, S 是任意 k 维空间. 令 $S'=\mathrm{Span}\{x_k,\cdots,x_n\}$. 则 $\dim S+\dim S'=k+(n-k+1)=n+1$. 所以 $\delta\ge 1$, 故 $\{x\neq 0\wedge x\in S\cap S'\}$ 非空. 这时有:

$$
\begin{aligned}
\sup_{\substack{x\neq 0\\ x\in S}}\frac{x^*Ax}{x^*x}\ge\sup_{\substack{x\neq 0\\ x\in S\cap S'}}\frac{x^*Ax}{x^*x}\ge\inf_{\substack{x\neq 0\\ x\in S\cap S'}}\frac{x^*Ax}{x^*x}\\
\ge\inf_{\substack{x\neq 0\\ x\in S'}}\frac{x^*Ax}{x^*x}\ge\min_{\substack{x\neq 0\\ x\in S'}}\frac{x^*Ax}{x^*x}=\lambda_k
\end{aligned}
$$

总结一下就是

$$\sup_{\substack{x\neq 0\\ x\in S}}\frac{x^*Ax}{x^*x}\ge\lambda_k,\qquad\forall S\subset\mathbb{C}^n$$

所以自然就有:

$$\inf_{\substack{S\subset\mathbb{C}^n\\ \dim S=k}}\sup_{\substack{x\neq 0\\ x\in S}}\frac{x^*Ax}{x^*x}\ge\lambda_k$$

下面来研究这个等号是否一定能够取到: 令这里的 $x_1,x_2,\cdots,x_n$ 分别是 $\lambda_1,\cdots,\lambda_n$ 对应的特征向量, 且相互正交. 所以取 $S=\mathrm{Span}\{x_1,\cdots,x_k\}$ 即可取到等号, 故

$$\inf_{\substack{S\subset\mathbb{C}^n\\ \dim S=k}}\sup_{\substack{x\neq 0\\ x\in S}}\frac{x^*Ax}{x^*x}=\lambda_k$$

□

1.1.4 柯西交错定理

Theorem 5 (Weyl Theorem)**.** *将 Hermitian 阵 $A,B,A+B$ 的特征值都从小到大排序, 然后我们有:*

$$\lambda_i(A+B)\le\lambda_{i+j}(A)+\lambda_{n-j}(B),\qquad {\scriptstyle j\in[0,i-1]\atop i\in[1,n]}\qquad(1)$$

对称的, 我们还有:

$$\lambda_{i-j+1}(A)+\lambda_j(B)\le\lambda_i(A+B),\qquad {\scriptstyle j\in[1,i]\atop i\in[1,n]}\qquad(2)$$

证明. 这个定理看上去十分复杂, 但是思路很好理解. 先来证明 (1), 设

$S_1=\{x_i,\cdots,x_n\}, \qquad (A+B)x_i=\lambda_i(A+B)x_i.$

$S_2=\{y_1,\cdots,y_{i+j}\}, \quad Ay_i=\lambda_i(A)y_i.$

$S_3=\{z_1,\cdots,z_{n-j}\}, \quad Bz_i=\lambda_i(B)z_i.$

很容易发现 $\delta+2n=\dim S_1+\dim S_2+\dim S_3=(i+j)+(n-j)+(n-i+1)=2n+1$. 故 $\delta\geq 1$, 所以 $S_1\cap S_2\cap S_3$ 非空, 故对 $x\in S_1\cap S_2\cap S_3$, 有:

$$\begin{aligned}\lambda_i(A+B)\underbrace{\leq}_{S_1} x^*(A+B)x &= x^*Ax+x^*Bx\\ &\leq \underbrace{\lambda_{i+j}(A)}_{S_2}+\underbrace{\lambda_{n-j}(B)}_{S_3}\end{aligned}$$

同理即可证明 (2). □

Corollary 1. *令 A,B 为 Hermitian 阵, 并设 B 有 π 个正特征值, ν 个负特征值. 则显然有 $\substack{\lambda_{n-\pi}(B)\leq 0\\ \lambda_{\nu+1}(B)\geq 0}$. 故:*

[1]

$$\begin{aligned}\lambda_i(A+B)&\leq\lambda_{i+\pi}(A)+\underbrace{\lambda_{n-\pi}(B)}_{\leq 0}\\ &\leq\lambda_{i+\pi}(A), \quad i\in[1,n-\pi]\end{aligned}$$

[2]

$$\begin{aligned}\lambda_{i-\nu}(A)+\underbrace{\lambda_{1+\nu}(B)}_{\geq 0}&\leq\lambda_i(A+B)\\ \lambda_{i-\nu}(A)&\leq\lambda_i(A+B), \quad i\in[\nu+1,n]\end{aligned}$$

Fact 10. *$\forall z\in\mathbb{C}^n$, 是非 0 向量, 且 $n\geq 2$, 则有: $\lambda_{n-1}(zz^*)=0=\lambda_2(zz^*)$*

proof. 考虑两个特征值对, $\substack{\lambda_1\neq 0,x\neq 0\\ \lambda_2\neq 0,y\neq 0}$. 因为 $zz^*x=\lambda_1 x$, 所以 $z^*x\neq 0$, 同理可得 $z^*y\neq 0$.

于是就有:

$$\begin{aligned}0\neq(z^*y)^*(z^*x)=y^*zz^*x=\lambda_1 y^*x\neq 0\\ 0\neq(z^*x)^*(z^*y)=x^*zz^*y=\lambda_2 x^*y\neq 0\end{aligned}$$

显然, $(z^*y)^*(z^*x)=\overline{(z^*x)^*(z^*y)}$, 所以 $\lambda_1 y^*x=\overline{\lambda_2 x^*y}$ 所以 $\lambda_1=\overline{\lambda_2}$. 又因为 zz^* 是 Hermitian 阵, 所以其特征值都是实数, 所以 $\lambda_1=\lambda_2$. □

Fact 11. *$\forall v\in\mathbb{C}^n$, 是非 0 向量, 且 $n\geq 2$, 则有 vv^* 的非 0 特征向量 >0.*

proof. 设 $\lambda\neq 0, x\neq 0$ 为 vv^* 的非 0 特征值对. 则,

$$0<\sum_{i=1}^n|[v^*x]_i|^2=(v^*x)^*(v^*x)=x^*vv^*x=\lambda x^*x=\lambda\sum_{i=1}^n|x_i|^2$$

. 故 $\lambda>0$. □

Theorem 6 (Cauchy's Interlacing Theorem)**.** *$n \geq 2$, A 是 Hermitian 阵, 那么: 对于 $B = A + vv^*$, 有:*

$$\lambda_1(A) \leq \lambda_1(B) \leq \cdots \leq \lambda_n(A) \leq \lambda_n(B)$$

proof. 因为 vv^* 有最多 1 个正特征值, 且其他的特征值都为 0, 所以:

$$\lambda_i(A + vv^*) \leq \lambda_{i+1}(A) + \underbrace{\lambda_{n-1}(vv^*)}_{\leq 0}$$

$$\lambda_i(B) \leq \lambda_{i+1}(A)$$

$$\lambda_{i-1}(A) + \underbrace{\lambda_{i+1}(vv^*)}_{\geq 0} \leq \lambda_i(A + vv^*)$$

$$\lambda_{i-1}(A) \leq \lambda_i(B)$$

□

1.1.5 文章中将会用到的一些结论

Define 4 (为正的特征值个数)**.** *文章中说的特征值个数, 是真的个数, 如果一个矩阵 $A \in \mathbb{R}^{n\times n}$, 有两个特征值, λ_1, λ_2, 这个时候, 是算两个特征值的.*

△ 注: 反正最有要保证 A 有 n 个特征值. i.e. 我们对于不平行的特征向量构成的特征值, 我们都算作不同的特征值.

Fact 12. *如果 $A \in \mathbb{R}^{n\times n}$ 最多只有一个正特征值, 则存在 $B \in \mathbb{R}^{n\times n}, v \in \mathbb{R}^n$, 使得 $A = B + vv^*$, 且 B 是一个半负定矩阵.*

proof. 根据 [Courant-Fischer Theorem] 我们知道, 存在一个 $x \in \mathbb{R}^n$, 使得 $x^*Ax = \lambda_{\max}x^*x$, 且 $x^*x = 1$.

这个时候, 我们去找一个 $v \in \mathbb{R}^n$, 使得 $v^*x = \lambda_{\max}^{1/2}$. (可以直接取 $v = \lambda_{\max}^{1/2}x$) 显然, 这样的 v 存在很多个, 我们任意取一个.

这个时候, 令 $B = A - vv^*$, 就有:

$$\begin{aligned} x^*Bx &= x^*Ax - x^*vv^*x \\ &= \lambda_{\max} - (v^*x)^*(v^*x) \\ &= \lambda_{\max} - \lambda_{\max}^{1/2}\lambda_{\max}^{1/2} \\ &= 0 \end{aligned}$$

因为 A 只有一个为正的特征向量, vv^* 的所有特征向量非负. 所以, 对于 $\forall y \in \mathbb{R}^n$ 且 y 与 x 不平行, 都有 $\substack{y^*Ay \leq 0 \\ y^*vv^*y = 0}$, 所以 $y^*By \leq 0$, 所以 B 是半负定矩阵. □

Define 5 (半定矩阵记法)**.** *文章中, 通常用 $A \preccurlyeq 0$ 表示 A 是负定的. 对称的, 一般用 $A \succcurlyeq 0$ 表示 A 是正定的.*

△ 注: $A \preccurlyeq B$ 表示 $A - B \preccurlyeq 0$.

Lemma (2.3)**.** *令 $A \in \mathbb{R}^{n\times n}$ 是一个对称阵. 令 $P \in \mathbb{R}^{m\times n}$. 如果 A 最多有一个正特征值, 则 PAP^T 最多有一个正特征值.*

proof. 由前面的结论知道, 我们可以令 $A = B + vv^*$, 其中 $B \preccurlyeq 0$. 然后就有:

$$\begin{aligned} PAP^T &= PBP^T + \underbrace{Pvv^*P^T}_{w:=Pv} \\ x^*PAP^Tx &= x^*PBP^Tx + xww^*x \\ &= (P^Tx)^*B(P^Tx) + xww^*x \end{aligned}$$

因为 $B \preccurlyeq 0$, 且 ww^* 最多只有一个正特征值, 所以 PAP^T 最多只有一个正特征值. □

Fact (2.4)**.** *令 $A \in \mathbb{R}^{n\times k}, B \in \mathbb{R}^{k\times n}$. AB 的非 0 特征值等于 BA 的非 0 特征值. 且具有相同的代数重数.*

proof.

$$\underbrace{\begin{bmatrix} I_{n\times n} & -A \\ 0 & I_{k\times k} \end{bmatrix}}_{C_1} \underbrace{\begin{bmatrix} AB & 0 \\ B & 0_{k\times k} \end{bmatrix}}_{S_1} \underbrace{\begin{bmatrix} I_{n\times n} & A \\ 0 & I_{k\times k} \end{bmatrix}}_{C_2} = \underbrace{\begin{bmatrix} 0_{n\times n} & 0 \\ B & BA \end{bmatrix}}_{S_2}$$

不难发现, $C_1C_2 = I$, 所以 S_1 和 S_2 相似. 所以 S_1, S_2 有相同的特征值.

考虑 S_1 的特征方程:

$$\begin{aligned} P(S_1) &= 0 \\ \det(S_1 - \lambda I) &= 0 \\ \det\left(\begin{bmatrix} AB - \lambda I_{n\times n} & 0 \\ B & -\lambda I_{k\times k} \end{bmatrix}\right) &= 0 \end{aligned}$$

计算行列式时, 从右下角按列展开, 不难发现, S_1 的特征值等于 AB 的特征值加上 k 个 0.

同理可得, S_2 的特征值等于 BA 的特征值加上 n 个 0.

又因为 S_1 和 S_2 的特征值相同, 所以不难发现, AB 和 BA 的非 0 特征值相同, 且具有相同的代数重数. □

Lemma (2.5)**.** *$A \in \mathbb{R}^{n\times n}$ 是一个对称阵, 且最多只有一个正特征值. 那么对任意 $B \in \mathbb{R}^{n\times n}, B \succcurlyeq 0$, 都有 BA 最多有一个正特征值.*

proof. 因为 B 半正定, 所以 B 实对称. 所以 $\exists P$ 使得 $B = P\Lambda P^T$ (Λ 为对角阵) 所以

$$\begin{aligned} B &= P\Lambda P^T \\ &= P\Lambda^{1/2}\Lambda^{1/2}P^T \\ &= (\Lambda^{1/2}P^T)^T(\Lambda^{1/2}P^T) \end{aligned}$$

所以, B 可以被写成 C^TC 的形式.

故:

$$\begin{aligned} BA &= C^TCA \\ &= C^T(CA) \end{aligned}$$

根据 Fact (2.4) 可以发现, $C^T(CA)$ 和 CAC^T 具有相同的非 0 特征值. 又因为 A 最多又一个正特征值, 所以 CAC^T 最多有一个正特征值, 所以 BA 最多有一个正特征值. □

Lemma (2.6)**.** 令 $A \in \mathbb{R}^{n\times n}$ 为一个实对称矩阵. 且 $A \geq 0$, A 最多有一个正特征值. 然后令 $w \in \mathbb{R}^n, w_i = \sum_{j=1}^{n} A_{ij}$ 且 $w_i > 0, \forall i$. 则

$$A \preccurlyeq \frac{ww^T}{\sum_i w_i}$$

proof. 令 $W = \operatorname{diag} w$. 根据 Lemma (2.3) $\mathcal{A} = W^{-1/2}AW^{-1/2}$ 最多只有一个正特征值. (这里 $\mathcal{A}$ 成立的原因是 W 中的元素都是正数) 然后令 $\sqrt{w} \in \mathbb{R}^n : \sqrt{w}(i) = \sqrt{w_i}$. 则

$$\begin{aligned}
\mathcal{A}\sqrt{w} &= W^{-1/2}AW^{-1/2}\sqrt{w} \\
&= W^{-1/2}A\,\mathbf{1} \\
&= W^{-1/2}w \\
&= \sqrt{w}
\end{aligned}$$

所以 $\sqrt{w}$ 取到了 $\mathcal{A}$ 唯一的正特征值, i.e. 1.

又有:

$$\sqrt{w}^T(\mathcal{A} - \frac{\sqrt{w}\sqrt{w}^T}{||\sqrt{}||^2})\sqrt{w} = \sqrt{w}^T\mathcal{A}\sqrt{w} - \sqrt{w}^T\sqrt{w} = 0$$

所以 $\mathcal{A} \preccurlyeq \frac{\sqrt{w}\sqrt{w}^T}{||\sqrt{w}||^2} = \frac{\sqrt{w}\sqrt{w}^T}{\sum_i w_i}$ [这里可以用和 Fact **??** 相同的语句来证明].

更进一步, 因为 $\mathcal{A} \preccurlyeq \frac{\sqrt{w}\sqrt{w}^T}{\sum_i w_i}$, 所以 $\forall x \in \mathbb{R}^n, x^Tx = 1$, 都有:

$$x^T(\mathcal{A} - \frac{\sqrt{w}\sqrt{w}^T}{\sum_i w_i})x \leq 0$$

因为 $W^{1/2}x \in \mathbb{R}^n$, 故

$$\begin{aligned}
(W^{1/2}x)^T(\mathcal{A} - \frac{\sqrt{w}\sqrt{w}^T}{\sum_i w_i})(W^{1/2}x) &\leq 0 \\
x^TW^{1/2}(\mathcal{A} - \frac{\sqrt{w}\sqrt{w}^T}{\sum_i w_i})W^{1/2}x &\leq 0 \\
x^T(W^{1/2}\mathcal{A}W^{1/2} - \frac{W^{1/2}\sqrt{w}\sqrt{w}^TW^{1/2}}{\sum_i w_i})W^{1/2}x &\leq 0 \\
x^T(A - \frac{ww^T}{\sum_i w_i})W^{1/2}x &\leq 0, \qquad \forall x \in \mathbb{R}^n
\end{aligned}$$

所以

$$A \preccurlyeq \frac{ww^T}{\sum_i w_i}$$

□

1.2 d-homogeneous multiaffine polynomial

△ 注: 这里的 d 不是在说变量的个数, 而是在说次数, 变量的个数一般设为 n.

Define 6 (log-concave polynomial)**.** 一个多项式 $p(x)$ 是 *log-concave* 的, 如果 $\log p(x)$ 是 *concave* 的.

d-homogeneous multiaffine polynomial 有一堆性质:

Fact 13. *如果 $p(x)$ 是 d-homogeneous multiaffine polynomial, 则*

$$d \cdot p(x) = \sum_{k=1}^{n} x_k \partial_k p(x)$$

△ 注: 不难发现, 这里的 $x_k \partial_k p(x)$ 相当于将多项式中所有包含 x_k 的项相加. 这样将所有变量 x_k 枚举完之后, 不难发现, $p(x)$ 中的每一项被加了 d 次. 上面的式子还可以写得更加简洁

$$d \cdot p = x^T (\nabla p)$$

Fact 14.

$$(d-1) \cdot \nabla p = \sum_{k=1}^{n} x_k \cdot \nabla(\partial_k p) = \nabla^2 p \cdot x$$

Fact 15.

$$(d-2) \cdot \nabla^2 p = \sum_{k=1}^{n} x_k \cdot \nabla^2(\partial_k p)$$

△ 注: 要理解上面的等式, 只需注意到 d-homogeneous multiaffine polynomial 求偏导之后, 仍然是 $(d-1)$-homogeneous multiaffine polynomial.

Fact 16.

$$\nabla^2 \log p = \frac{p \cdot \nabla^2 p - (\nabla p)(\nabla p)^T}{p^2}$$

proof.

$$\begin{aligned} \partial_i \partial_j p &= \partial_j (\frac{\partial_i p}{p}) \\ &= \frac{\partial_i \partial_j p \cdot p - \partial_i p \cdot \partial_j p}{p^2} \end{aligned}$$

□

△ 注: 因为 p 是 log-concave 的, 所以 $\log p$ 是 concave 的, 也就是说

$$\nabla^2 \log p \preccurlyeq 0$$

不妨假设 p 的所有变量都取正数, 就有:

$$\nabla^2 p \preccurlyeq \frac{(\nabla p)(\nabla p)^T}{p}$$

根据我们之前的结论, $vv^T, \forall v \in \mathbb{R}^n$ 最多有一个正特征值, 所以 $\nabla^2 p$ 最多有一个正特征值.

1.3 Matroid

比较熟悉, 略.

1.4 Simplicial Complexes

2 正文