奇异值分解（SVD，singular value decomposition）

Exercise 1. 对任何矩阵 $A=(a_{ij})$（即 A 的 (i,j) 元为 a_{ij}），证明 $\sum a_{ij}^2=\sum \sigma_i^2$，其中 $\sigma_1,\sigma_2,\dots$ 为 A 的所有奇异值。

证明. 假设 A 为 m 行 n 列，其中 $m\geq n$，若 $m<n$，则取其转置即可. 做奇异值分解，我们有

$$A_{m\times n}=U_{m\times n}\Lambda_{n\times n}V_{n\times n}$$

其中 V 是一个正交方阵，我们有 $V^\top V=VV^\top=E$, E 是一个单位方阵. 因此

$$\begin{aligned}\sum a_{ij}^2 =&\mathrm{tr}(A^\top A)\\ =&\mathrm{tr}(V^\top\Lambda U^\top U\Lambda V)\\ =&\mathrm{tr}(V^\top\Lambda^2 V)\\ =&\mathrm{tr}(VV^\top\Lambda^2)\\ =&\mathrm{tr}(\Lambda^2)\\ =&\sum\sigma_i^2\end{aligned}$$

□

Exercise 2. 说明正定矩阵 A 的奇异值分解与谱分解是一致的（因此 *Eckart-Young-Mirsky* 奇异值最佳逼近定理也适用于主成分，参见 *lecture15*，*P16* 定理 *2* 以及 *lecture 13*，*P2*）。

证明. 记 $A_{n\times n}$ 的奇异值分解为 $A=P\Sigma Q$, 则我们有 $A^2=A^\top A=Q^\top\Sigma^2Q$, 这表明 Σ^2 的对角元是 A^2 的特征值. 另一方面，我们又有 A 的谱分解为 $A=U\Lambda U^\top$, 因此 $A^2=U\Lambda^2U^\top$. 从而 Λ^2 的对角元与 Σ^2 的对角元都是 A^2 的特征根，因此他们对应相等. 从而我们有 $\lambda_i^2=\sigma_i^2$, 又由于 σ_i 与 λ_i 都是大于 0 的实数，因此 $\lambda_i=\sigma_i,\forall i=1,...,n$. 因此 A 的奇异值分解和谱分解是一致的. □

Exercise 3. 假设 $n\times 2$ 数据矩阵 $X=(\mathbf{x}_1,\mathbf{x}_2)$，$\|\mathbf{x}_1\|=a,\|\mathbf{x}_2\|=b,a>b$。假设 $\mathbf{x}_1^\top\mathbf{x}_2=0$，求 X 的奇异值分解。

证明. 其一个奇异值分解为

$$X=\left[\frac{\mathbf{x}_1}{a},\frac{\mathbf{x}_2}{b}\right]\begin{bmatrix}a&0\\0&b\end{bmatrix}$$

其中 $U=\left[\frac{\mathbf{x}_1}{a},\frac{\mathbf{x}_2}{b}\right]$, $V=E_2$ 为二阶单位矩阵 □

Exercise 4. 假设 A 是 $n\times m$ 矩阵，$\mathbf{u}\in\mathbb{R}^n,\mathbf{v}\in\mathbb{R}^m$，假设 $A\mathbf{v}=\alpha\mathbf{u},A^\top\mathbf{u}=\beta\mathbf{v}$，其中实数 $\alpha,\beta\neq 0$。证明 α 和 β 同号且 $\sqrt{\alpha\beta}$ 是 A 的一个奇异值。

证明. 由于 $A^\top A\mathbf{v}=A^\top\alpha\mathbf{u}=\alpha\beta\mathbf{v}$, 从而 $\mathbf{v}$ 是 $A^\top A$ 的一个特征向量, 对应的特征值为 $\alpha\beta$, 又因为 $A^\top A$ 是半正定的矩阵，从而 $\alpha\beta\geq 0$. 因为 $\alpha,\beta\neq 0$, 从而 α 与 β 是同号的. 因为 $\alpha\beta$ 是 $A^\top A$ 的一个特征值, 根据奇异值的定义, $\sqrt{\alpha\beta}$ 就是 A 的一个奇异值. □

Exercise 5. *(普氏分析)* 假设 X, Y 都是 $n \times p$ 矩阵，假设 $p \times p$ 方阵 $X^\top Y$ 的奇异值分解

$$X^\top Y = UDV^\top,$$

其中 U, V 都是 p 阶正交矩阵。证明 $Q = UV^\top$ 使得 $\|X - YQ\|_F^2$ 达到最小。

证明. 由于 Q 是一个正交矩阵，我们有:

$$\begin{aligned}
&||X - YQ||_F^2 \\
=&\mathrm{tr}(X^\top X) + \mathrm{tr}(Y^\top Y Q^\top Q) - 2\mathrm{tr}(X^\top Y Q^\top) \\
=&\mathrm{tr}(X^\top X) + \mathrm{tr}(Y^\top Y) - 2\mathrm{tr}(X^\top Y Q^\top)
\end{aligned}$$

因此问题转化为最大化 $\mathrm{tr}(X^\top Y Q^\top)$. 由奇异值分解,我们有 $\mathrm{tr}(X^\top Y Q^\top) = \mathrm{tr}(UDV^\top Q^\top) = \mathrm{tr}(DV^\top Q^\top U) = \sum_i D_{ii}(V^\top Q^\top U)_{ii}$. 记 $\mathbf{v}_i$ 和 $\mathbf{u}_i$ 分别为 V 和 U 的第 i 列, 我们有 $(V^\top Q^\top U)_{ii} = \mathbf{v}_i^\top Q^\top \mathbf{u}_i$, 从而 $(V^\top Q^\top U)_{ii} \le ||Q\mathbf{v}_i||_2 ||\mathbf{u}_i||_2 = 1$, 因此当 $Q = UV^\top$ 时，$(V^\top Q^\top U)_{ii}$ 才会取到最大值为 1. □

典则相关分析 (CCA，canonical correlation analysis)：

Exercise 6. 假设 y 是随机变量，x 是 $p \times 1$ 随机向量，假设它们的协方差矩阵为

$$\Sigma_{cov}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} > 0,$$

其中 $\Sigma_{yy} = \sigma_y^2$ 是标量（正实数），Σ_{xy} 是 $p \times 1$ 向量。

(a) 试求 x, y 的第一典则相关系数 $\sqrt{\lambda_1}$ *(也是唯一的非 0 典则相关系数)* 和第一对典则变量 (ξ_1, η_1)。

(b) 假设 y, x 均值都为 *0*，满足线性模型

$$y = \beta^\top x + \epsilon, \epsilon \sim (0, \sigma^2), \epsilon \perp\!\!\!\perp x,$$

证明该模型等价于

$$\eta_1 = \sqrt{\lambda_1}\xi_1 + \delta_1, \delta_1 \sim (0, 1 - \lambda_1), \delta_1 \perp\!\!\!\perp \xi_1$$

(验证：$\beta^\top x = \sqrt{\lambda_1}\sigma_y \times \xi_1, y = \sigma_y \times \eta_1$)。

证明. (a)

根据第一典则相关系数的定义，需要寻找 $\boldsymbol{\alpha}$ 和 β 来使得 $\mathrm{Cov}(\boldsymbol{\alpha}^\top \mathbf{x}, by)$ 最大, 即最大化

$$\frac{\boldsymbol{\alpha}^\top \Sigma_{\mathrm{xy}} b}{\sqrt{\boldsymbol{\alpha}^\top \Sigma_{\boldsymbol{xx}} \boldsymbol{\alpha}}\sqrt{b\sigma_y^2 b}}$$

为了使得 Cauchy-Schwarz 不等式，能够取得最大，此时第一典则相关系数为 $\sqrt{\lambda_1} = \frac{\sqrt{\Sigma_{y\boldsymbol{x}}\Sigma_{\boldsymbol{xx}}^{-1}\Sigma_{\boldsymbol{x}y}}}{\sigma_y}$, 第一对典则变量为 $\xi_1 = \frac{\Sigma_{y\boldsymbol{x}}\Sigma_{\boldsymbol{xx}}^{-1}\boldsymbol{x}}{\sqrt{\Sigma_{y\boldsymbol{x}}\Sigma_{\boldsymbol{xx}}^{-1}\Sigma_{\boldsymbol{x}y}}}$, $\eta_1 = \frac{y}{\sigma_y}$.

(b) 对于式子 $y = \beta^\top \mathbf{x} + \epsilon$, 左右两边同时对 $\mathbf{x}$ 取相关系数, 我们就可以得到 $\boldsymbol{\beta} = \Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathrm{xy}}$, 从而

$$\boldsymbol{\beta}^\top \mathbf{x} = \Sigma_{y\mathrm{x}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x} = \frac{\Sigma_{y\mathrm{x}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}}{\sqrt{\Sigma_{y\boldsymbol{x}}\Sigma_{\boldsymbol{xx}}^{-1}\Sigma_{\boldsymbol{x}y}}}\frac{\sqrt{\Sigma_{y\boldsymbol{x}}\Sigma_{\boldsymbol{xx}}^{-1}\Sigma_{\boldsymbol{x}y}}}{\sigma_y}\sigma_y = \sqrt{\lambda_1}\sigma_y \times \xi_1.$$

由于 $\eta_1 = \frac{y}{\sigma_y}$, 从而 $y = \sigma_y \times \eta_1$. 又由于 $\sigma_y^2 = \Sigma_{yy} = \sqrt{\Sigma_{y\boldsymbol{x}}\Sigma_{\boldsymbol{xx}}^{-1}\Sigma_{\boldsymbol{x}y}} + \sigma^2$, 从而可以得到 $\lambda_1\sigma_y^2 + \sigma^2 = \sigma_y^2$, 即 $\delta_1 = \frac{\epsilon}{\sigma_y} \sim (0, 1 - \lambda_1)$ □

Exercise 7. 假设随机向量 $\begin{pmatrix} y_{q\times 1} \\ x_{p\times 1} \end{pmatrix}$ 的协方差矩阵为相关系数矩阵 $R=(\rho_{ij})$，其中 $\rho_{ii}=1, i=1,...,p+q$。证明 x 和 y 的第一典则相关系数 $\sqrt{\lambda_1}\geq \rho_{ij}, i=1,...,q; j=q+1,...,q+p$。

证明. 根据第一典则相关系数的定义, $\sqrt{\lambda_1}=\max_{\mathbf{a}^\top\Sigma_{\mathbf{xx}}\mathbf{a}=\mathbf{b}^\top\Sigma_{\mathbf{yy}}\mathbf{b}=1}\mathrm{Cov}(\mathbf{a}^\top\mathbf{x},\mathbf{b}^\top\mathbf{y})=\max_{\mathbf{a}^\top\Sigma_{\mathbf{xx}}\mathbf{a}=\mathbf{b}^\top\Sigma_{\mathbf{yy}}\mathbf{b}=1}\mathbf{a}^\top\Sigma_{\mathbf{xy}}\mathbf{b}$, 由于 $\Sigma_{\mathbf{xx}}=E_{p\times p}$, $\Sigma_{\mathbf{yy}}=E_{q\times q}$, 从而 $||\mathbf{a}||_2=||\mathbf{b}||_2=1$. 因此约束优化问题为

$$\sqrt{\lambda_1}=\max_{||\mathbf{a}||_2=||\mathbf{b}||_2=1}\mathbf{a}^\top\Sigma_{\mathbf{xy}}\mathbf{b}$$

当取 $\mathbf{a}$ 的第 j 个分量为 1, 其余分量为 0, 取 $\mathbf{b}$ 的第 i 个分量为 1, 其余分量为 0 的时候, 右边会等于 ρ_{ij}. 从而得证. □

Exercise 8. 假设随机向量 $\begin{pmatrix} y_{q\times 1} \\ x_{p\times 1} \end{pmatrix}$ 的协方差矩阵为 $\Sigma=(1-\rho)I_m+\rho\mathbf{1}_m\mathbf{1}_m^\top, m=p+q, \rho>-1/(m-1)$，试求 x,y 的第一典则相关系数和典则变量（其中 I_m 是 m 阶单位阵，$\mathbf{1}_m$ 是长度为 m 的元素全为 *1* 的向量）。

证明. 按照定义, 为了求解第一典则相关系数, 我们要求解以下的最大化问题,

$$\max_{\boldsymbol{a}^\top\Sigma_{\boldsymbol{xx}}\boldsymbol{a}=\boldsymbol{b}^\top\Sigma_{\boldsymbol{yy}}\boldsymbol{b}=1}\boldsymbol{a}^\top\Sigma_{\boldsymbol{xy}}\boldsymbol{b}$$

由于 $\Sigma_{\boldsymbol{xx}}=(1-\rho)I_p+\rho\mathbf{1}_p\mathbf{1}_p^\top$, $\Sigma_{\boldsymbol{yy}}=(1-\rho)I_q+\rho\mathbf{1}_q\mathbf{1}_q^\top$, $\Sigma_{\boldsymbol{xy}}=\rho\mathbf{1}_p\mathbf{1}_q^\top$, 带入上式可得优化问题为,

$$\begin{aligned}&\max\rho(\mathbf{1}_p^\top\boldsymbol{a})(\mathbf{1}_q^\top\boldsymbol{b})\\ \text{subject to:}&(1-\rho)\boldsymbol{a}^\top\boldsymbol{a}+\rho(\mathbf{1}_p^\top\boldsymbol{a})^2=1\\ &(1-\rho)\boldsymbol{b}^\top\boldsymbol{b}+\rho(\mathbf{1}_q^\top\boldsymbol{b})^2=1\end{aligned}$$

利用拉格朗日乘子法, 拉格朗日函数为

$$L(\boldsymbol{a},\boldsymbol{b},\lambda,\mu)=\rho(\mathbf{1}_p^\top\boldsymbol{a})(\mathbf{1}_q^\top\boldsymbol{b})+\lambda\big[(1-\rho)\boldsymbol{a}^\top\boldsymbol{a}+\rho(\mathbf{1}_p^\top\boldsymbol{a})^2-1\big]+\mu\big[(1-\rho)\boldsymbol{b}^\top\boldsymbol{b}+\rho(\mathbf{1}_q^\top\boldsymbol{b})^2-1\big]$$

其一阶静止条件为,

$$\frac{\partial L}{\partial\boldsymbol{a}}=\rho(\mathbf{1}_q^\top\boldsymbol{b})\mathbf{1}_p+\lambda\big[2(1-\rho)\boldsymbol{a}+2\rho(\mathbf{1}_p^\top\boldsymbol{a})\mathbf{1}_p\big]=0 \tag{1}$$

$$\frac{\partial L}{\partial\boldsymbol{b}}=\rho(\mathbf{1}_p^\top\boldsymbol{a})\mathbf{1}_q+\mu\big[2(1-\rho)\boldsymbol{b}+2\rho(\mathbf{1}_q^\top\boldsymbol{b})\mathbf{1}_q\big]=0 \tag{2}$$

$$\frac{\partial L}{\partial\lambda}=(1-\rho)\boldsymbol{a}^\top\boldsymbol{a}+\rho(\mathbf{1}_p^\top\boldsymbol{a})^2-1=0 \tag{3}$$

$$\frac{\partial L}{\partial\mu}=(1-\rho)\boldsymbol{b}^\top\boldsymbol{b}+\rho(\mathbf{1}_q^\top\boldsymbol{b})^2-1=0 \tag{4}$$

$$\tag{5}$$

对 (1) 求和得到

$$p\rho(\mathbf{1}_q^\top\boldsymbol{b})+\lambda\big[2(1-\rho)(\mathbf{1}_p^\top\boldsymbol{a})+2p\rho(\mathbf{1}_p^\top\boldsymbol{a})\big]=0$$

从而可以得到

$$\mathbf{1}_p^\top\boldsymbol{a}=\frac{-p\rho(\mathbf{1}_q^\top\boldsymbol{b})}{2\lambda(1-\rho+p\rho)} \tag{6}$$

将 (6) 代入 (1) 我们可以发现,

$$\rho(\mathbf{1}_q^\top \boldsymbol{b})\mathbf{1}_p + \lambda\big[2(1-\rho)\boldsymbol{a} + 2\rho\frac{-p\rho(\mathbf{1}_q^\top \boldsymbol{b})}{2\lambda(1-\rho+p\rho)}\mathbf{1}_p\big] = 0$$

这意味着 $\boldsymbol{a}$ 的各个分量全都相等, 同样的逻辑我们可以得到 $\boldsymbol{b}$ 的各个分量也全都相等, 记 $\boldsymbol{a} = a\mathbf{1}_p, \boldsymbol{b} = b\mathbf{1}_q$, 从而利用式 (3) 和 (4), 我们可以得到

$$(1-\rho)pa^2 + \rho p^2 a^2 = 1$$
$$(1-\rho)qb^2 + \rho q^2 b^2 = 1$$

解得 $a = \pm\frac{1}{\sqrt{p(1-\rho+\rho p)}}$, $b = \pm\frac{1}{\sqrt{q(1-\rho+\rho q)}}$, 为了使得该最大化问题的目标函数最大, $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 应该同号, 从而该最大化问题的解为

$$(\boldsymbol{a}, \boldsymbol{b}) = (\frac{1}{\sqrt{p(1-\rho+\rho p)}}\mathbf{1}_p, \frac{1}{\sqrt{q(1-\rho+\rho q)}}\mathbf{1}_q)$$

或者

$$(\boldsymbol{a}, \boldsymbol{b}) = (-\frac{1}{\sqrt{p(1-\rho+\rho p)}}\mathbf{1}_p, -\frac{1}{\sqrt{q(1-\rho+\rho q)}}\mathbf{1}_q)$$

代入 $\boldsymbol{a}$ 和 $\boldsymbol{b}$, 得到最大值为

$$\frac{\rho\sqrt{pq}}{\sqrt{(1-\rho+\rho p)(1-\rho+\rho q)}}$$

此为第一典则相关系数. 对应的第一典则变量为,

$$(\boldsymbol{\xi}_1, \boldsymbol{\eta}_1) = (\boldsymbol{a}^\top \mathbf{x}, \boldsymbol{b}^\top \mathbf{y}) = (\pm\frac{1}{\sqrt{p(1-\rho+\rho p)}}\sum_{i=1}^p x_i, \pm\frac{1}{\sqrt{q(1-\rho+\rho q)}}\sum_{j=1}^q y_j)$$

□

相似度/距离，配列 (seriation)，多维标度法（MDS）

Exercise 9. *证明在 p 维欧氏空间中至多存在 $p+1$ 个点，它们两两之间的欧氏距离全相等。*

证明. 记 $\mathbf{x}_1, ..., \mathbf{x}_{p+2} \in \mathbb{R}^p$, 满足 $||\mathbf{x}_i - \mathbf{x}_j|| = d, \forall i \neq j$. 令 $\mathbf{y}_i = \mathbf{x}_i - \mathbf{x}_1, \forall i = 1, ..., p+2$, 则我们仍有 $||\mathbf{y}_i - \mathbf{y}_j|| = d, \forall i \neq j$, 同时我们还有 $||\mathbf{y}_i|| = d, \forall i = 2, ..., p+2$, 与 $\mathbf{y}_1 = \mathbf{0}$. 记 $Y = \begin{bmatrix}\mathbf{y}_1 & \dots & \mathbf{y}_{p+2}\end{bmatrix}_{p\times(p+2)}$. 则 $G = Y^\top Y$ 是一个秩至多为 p 的对称矩阵. 计算可以发现 $\mathbf{y}_i^\top \mathbf{y}_j = -\frac{||\mathbf{y}_i - \mathbf{y}_j||_2^2 - ||\mathbf{y}_i||_2^2 - ||\mathbf{y}_j||_2^2}{2} = \frac{d^2}{2}, \forall i \neq j, i, j \geq 2$. 因此我们可以得到,

$$G = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{G} \end{bmatrix}$$

而 $\tilde{G} = \frac{d^2}{2}\mathbf{I}_{p+1} + \frac{d^2}{2}\mathbf{1}_{p+1}\mathbf{1}_{p+1}^\top$ 是一个秩为 $p+1$ 的矩阵, 这迫使 G 的秩也为 $p+1$, 这与 $G = Y^\top Y$ 是相矛盾的. □

Exercise 10. *假设 $S = (s_{ij})$ 是 $n \times n$ 对称矩阵，$s_{ij} \geq 0$，记 $d_i = \sum_{j=1}^n s_{ij}, D = diag(d_1, ..., d_n)$，假设 $d_i > 0, i = 1, ..., n$。令拉普拉斯矩阵（Laplacian）$L \triangleq D - S$，令 $B = D^{-1}S$。*

(a) 证明 L 是双向中心化的（即行和、列和都为 0），且 0 一定是 L 的最小特征根。

(b) 对任何 $x = (x_1, ..., x_n)^T \in \mathbb{R}^n$，

$$x^T L x = \frac{1}{2} \sum_{1 \le i,j \le n} s_{ij}(x_i - x_j)^2.$$

(c) 说明 B 的所有特征根 λ 都是实数，证明 $|\lambda| \le 1$，且 *1* 是最大特征根。

(d) (选) 若 $s_{ij} > 0$，我们称节点 i, j 是相邻的，记作 $i \sim j$。证明 $\lambda = -1$ 是 B 的特征根当且仅当图所有 n 个节点可由两种颜色染色，使得相邻的节点不同色。

证明. (a)

$$L = (l_{ij}) = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} - \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ & \vdots & \\ s_{n1} & \cdots & s_{nn} \end{pmatrix}$$

$$\sum_{i=1}^n l_{ij} = d_j - \sum_{i=1}^n s_{ij} = 0$$

$$\sum_{j=1}^n l_{ij} = d_i - \sum_{j=1}^n s_{ij} = 0$$

所以 L 是双向中心化的。由于 L 双向中心化，可知

$$L1 = D1 - S1 = 0$$

0 为 L 的特征根，再由 (b) 知 L 半正定，所以 0 为 L 的最小特征根。

(b)

$\forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n\times 1}$，我们有

$$\begin{aligned}
& x^\top L x \\
=& x^\top (D - S) x \\
=& \sum_{i=1}^n d_i x_i^2 - \left(\sum_{i=1}^n x_i s_{i1}, \ldots, \sum_{i=1}^n x_i s_{in} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
=& \sum_{i=1}^n \sum_{j=1}^n s_{ij} x_i^2 - \sum_{j=1}^n \sum_{i=1}^n x_j x_i s_{ij} \\
=& \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n s_{ij} (x_i - x_j)^2
\end{aligned}$$

(c)

因为对任何方阵 A, B, AB 和 BA 有相同的非 0 特征根，故 $B = D^{-1}S$ 与 $D^{-1/2} S D^{-1/2}$ 有相同的非 0 特征根，后者是对称矩阵，特征根为实数，所以 B 的特征根全是实数。假设 λ 是 B 的一个特

征根，对应的特征向量为 $x = (x_1, ..., x_n)^T \neq 0$，则有特征方程

$$Bx = D^{-1}Sx = \lambda x,$$

记 $|x_k| = \max_{1 \leq i \leq n} |x_i|$，必定 $x_k \neq 0$。考察特征方程两边的第 k 个分量

$$|\lambda x_k| = \left| \sum_{j=1}^{n} \frac{s_{kj} x_j}{d_k} \right| \leq \sum_{j=1}^{n} \frac{s_{kj} |x_j|}{d_k} \leq \sum_{j=1}^{n} \frac{s_{kj} |x_k|}{d_k} = |x_k|$$

由此得 $|\lambda| \leq 1$ (因为 $x_k \neq 0$)。另外，由 (a)，$D\mathbf{1} = S\mathbf{1} \Rightarrow D^{-1}S\mathbf{1} = \mathbf{1}$，这表明 1 是 B 的最大的特征根。

(d) 如果 $\lambda = -1$ 是 B 的特征根，我们记其对应的特征向量为 $\mathbf{v}$, 则我们有

$$B\mathbf{v} = \lambda \mathbf{v}$$

即

$$S\mathbf{v} = -D\mathbf{v}$$

从而

$$(S + D)\mathbf{v} = \mathbf{0}$$

因此 $\forall i, \sum_{j=1}^{n} s_{ij}(v_i + v_j) = 0$, 若 $s_{ij} > 0$, 则必须有 $v_i = -v_j$, 这意味着节点 i 和 j 一旦相邻，那么其颜色必不相同. 在图是全联通的情况下，整个图可以由两种颜色来染色.

如果整个图可以用两种颜色染色，相邻的节点颜色不同，记其中一种颜色的全体节点为 A, 另一种颜色的全体节点为 B. 我们定义如下向量 $v_i = \begin{cases} 1 & i \in A \\ -1 & i \in B \end{cases}$, 则

$$(B\mathbf{v})_i = \sum_{j=1}^{n} \frac{s_{ij}}{d_i} v_j$$

如果 $s_{ij} > 0$, 则说明 i 和 j 是相邻的, 此时按照我们对 $\mathbf{v}$ 的定义, 必须有 $v_i = -v_j$. 因此上式可以写为

$$(B\mathbf{v})_i = \sum_{j=1}^{n} \frac{s_{ij}}{d_i} v_j = -v_i \sum_{j=1}^{n} \frac{s_{ij}}{d_i} = -v_i$$

从而我们得到了 $B\mathbf{v} = -\mathbf{v}$. 因此 -1 是 B 的一个特征根.

$\square$