概念解析：瑞利商

瑞利商（Rayleigh quotient）的概念在线性判别分析（LDA）、主成分分析（PCA）、谱聚类（Spectral clustering）等算法中均有涉及。

注：有些资料将“Rayleigh quotient”翻译为“瑞利熵”，“quotient”的本意是商的意思，而“entropy”才是信息论中的熵，所以个人认为翻译成“瑞利商”更合适

1. 瑞利商

给定一个Hermite矩阵A和非零向量x，**瑞利商**$R(A,x)$定义为：

$$R(A,x)=\frac{x^*Ax}{x^*x} \tag{1}$$

式（1）是两个二次型相除，因此对于向量x具有缩放不变性。即设c为一常数，则对于$x\longrightarrow cx$，有

$$R(A,cx)=\frac{(cx)^*Acx}{(cx)^*cx}=\frac{c^*c}{c^*c}\frac{x^*Ax}{x^*x}=R(A,x)$$

因此，不失一般性地，我们可以令$\|x\|^2=x^Tx=1$，然后以此为约束，考虑寻找函数$R(A,x)=x^TMx$的驻点。

根据拉格朗日乘子法，有

$$L(x)=x^TAx-\lambda(x^Tx-1)$$

其中λ为拉格朗日乘子。求导，得

$$\begin{aligned}&\frac{dL(x)}{dx}=0\\&\Rightarrow 2x^TA-2\lambda x^T=0\\&\Rightarrow 2Ax-2\lambda x=0\\&\Rightarrow Ax=\lambda x\end{aligned}$$

所以

$$R(A,x)=\frac{x^TAx}{x^Tx}=\lambda\frac{x^Tx}{x^Tx}=\lambda$$

由此可以导出瑞利商的一条**重要性质**：

- 瑞利商$R(A,x)$的极值就是矩阵A的特征值，其最大值为A的最大特征值λ_{max}，最小值为A的最小特征值λ_{min}；$R(A,x)$的极值点对应A的特征向量。

这条性质是主成分分析和典型相关性分析等方法的基础。

2. 广义瑞利商

广义瑞利商$R(A,B,x)$的定义为：

$$R(A,B,x)=\frac{x^*Ax}{x^*Bx} \tag{2}$$

其中A,B均为Hermite矩阵，B为正定阵，x是非零向量。

同理可以得到：

$$L(x)=x^*Ax-\lambda(x^*Bx-1)$$

令导数为零，有

$$\frac{dL(x)}{dx}=0\Rightarrow Ax=\lambda Bx$$

我们令$x=B^{-\frac{1}{2}}z$，这一步又被称为标准化，得到

$$AB^{-\frac{1}{2}}z=\lambda B^{\frac{1}{2}}z\Rightarrow B^{-\frac{1}{2}}AB^{-\frac{1}{2}}z=\lambda z$$

于是式（2）就可以化为普通瑞利商的形式：

$$R(A,B,x)=\frac{z^*B^{-\frac{1}{2}}AB^{-\frac{1}{2}}z}{z^*z} \tag{3}$$

由于

$$Ax=\lambda Bx\Rightarrow B^{-1}Ax=\lambda x$$

所以$B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$和$B^{-1}A$具有相同的特征值（以及不同的特征向量）。

于是求解式（3）的广义瑞利商的最大值和最小值就转变为求解矩阵$B^{-1}A$特征值的最大值和最小值。