回归分析 (01714601)
课程主页：http://staff.ustc.edu.cn/~ynyang/lm2020

第十讲 欧氏空间中的投影

2020.3.20

$$P_X = X(X^\top X)^{-1}X^\top$$

预备：向量之间的投影或正交化

参考：James H Stapleton (1995)Linear statistical models。Wiley

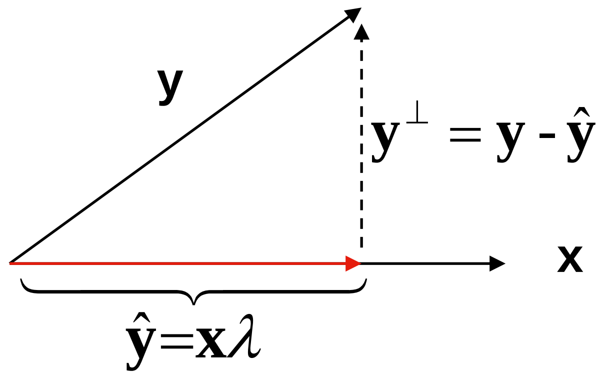
首先考察向量之间的投影

向量之间的投影：向量$\mathbf{y}\in R^n$在向量$\mathbf{x}\in R^n$上的投影$\hat{\mathbf{y}}$ 满足

（i）存在实数λ，$\hat{\mathbf{y}}=\mathbf{x}\lambda$

（ii）$\mathbf{y}-\hat{\mathbf{y}}\perp\mathbf{x}$，即$\mathbf{x}^{\mathrm{T}}\left(\mathbf{y}-\hat{\mathbf{y}}\right)=(\mathbf{x},\mathbf{y}-\mathbf{x}\lambda)=0$

向量的数乘：
向量在左,标量在右



由（ii），$\mathbf{x}^{\mathsf T}\left(\mathbf{y}-\hat{\mathbf{y}}\right)=0 \Rightarrow \mathbf{x}^{\mathsf T}\mathbf{y}=\mathbf{x}^{\mathsf T}\mathbf{x}\lambda \Rightarrow \quad \lambda=(\mathbf{x}^{\mathsf T}\mathbf{x})^{-1}\mathbf{x}^{\mathsf T}\mathbf{y}=\dfrac{\mathbf{x}^{\mathsf T}\mathbf{y}}{\mathbf{x}^{\mathsf T}\mathbf{x}}$

$$\Rightarrow \quad \hat{\mathbf{y}}=\mathbf{x}\lambda=\mathbf{x}(\mathbf{x}^{\mathsf T}\mathbf{x})^{-1}\mathbf{x}^{\mathsf T}\mathbf{y}=\left\{\mathbf{x}(\mathbf{x}^{\mathsf T}\mathbf{x})^{-1}\mathbf{x}^{\mathsf T}\right\}\mathbf{y}\triangleq P_{\mathbf{x}}\mathbf{y},$$

其中$P_{\mathbf{x}}=\mathbf{x}(\mathbf{x}^{\mathsf T}\mathbf{x})^{-1}\mathbf{x}^{\mathsf T}=\mathbf{x}\mathbf{x}^{\mathsf T}/\mathbf{x}^{\mathsf T}\mathbf{x}$称为是$\mathbf{x}$对应的投影矩阵。

$\Rightarrow \mathbf{y}^{\perp}\triangleq\mathbf{y}-\hat{\mathbf{y}}=\mathbf{y}-P_{\mathbf{x}}\mathbf{y}=$（$I-P_{\mathbf{x}}$）$\mathbf{y}$。

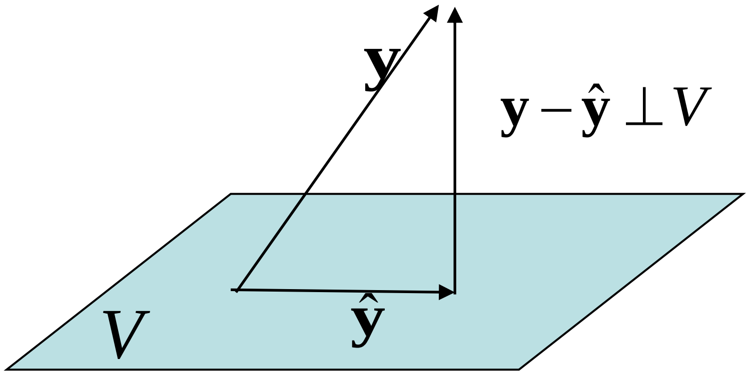
$\mathbf{y}^{\perp}$与$\mathbf{x}$正交$\Rightarrow\|\mathbf{y}\|^2=\|\hat{\mathbf{y}}+\mathbf{y}^{\perp}\|^2=\|\hat{\mathbf{y}}\|^2+\|\mathbf{y}^{\perp}\|^2+2\hat{\mathbf{y}}^T\mathbf{y}^{\perp}=\|\hat{\mathbf{y}}\|^2+\|\mathbf{y}^{\perp}\|^2$

正交分解：$\mathbf{y}=\hat{\mathbf{y}}+\mathbf{y}^{\perp}$，　$\hat{\mathbf{y}}=P_{\mathbf{x}}\mathbf{y}$，　$\mathbf{y}^{\perp}=(I-P_{\mathbf{x}})\mathbf{y}$

勾股定理：$\|\mathbf{y}\|^2=\|\hat{\mathbf{y}}\|^2+\|\mathbf{y}^{\perp}\|^2$

欧氏空间的投影

定义. 向量$\mathbf{y}\in R^n$在子空间$V\subset R^n$上的正交投影 定义为 $\hat{\mathbf{y}}=P(\mathbf{y}\,|\,V)\in V$, 使得$\mathbf{y}-\hat{\mathbf{y}}\perp V$.



正交分解： $\mathbf{y}=\hat{\mathbf{y}}+\mathbf{y}^{\perp}$, 其中$\hat{\mathbf{y}}\perp\mathbf{y}^{\perp}$

平方和分解: $\|\mathbf{y}\|^2=\|\hat{\mathbf{y}}\|^2+\|\mathbf{y}^{\perp}\|^2$

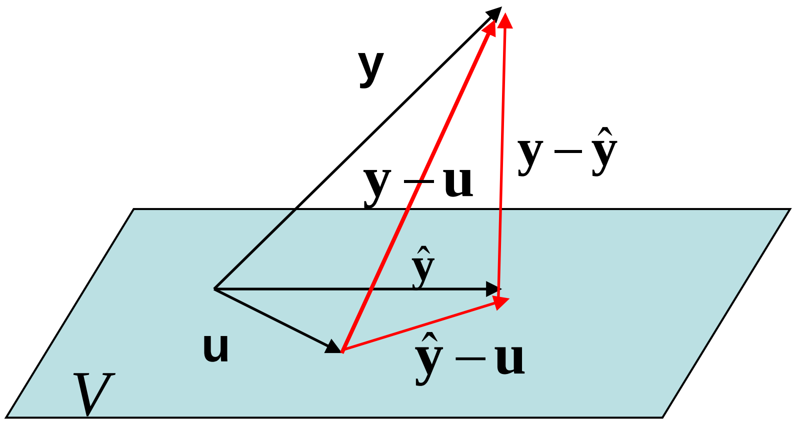
投影唯一：假设$\hat{\mathbf{y}}_1,\hat{\mathbf{y}}_2$都是$\mathbf{y}$的投影，$\mathbf{y}-\hat{\mathbf{y}}_i\perp V$, 特别地$\mathbf{y}-\hat{\mathbf{y}}_i\perp\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2$,

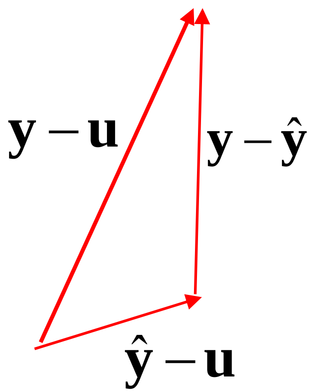
$\Rightarrow\left(\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2,\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2\right)=\left(\mathbf{y}-\hat{\mathbf{y}}_2-(\mathbf{y}-\hat{\mathbf{y}}_1),\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2\right)$

$=\left(\mathbf{y}-\hat{\mathbf{y}}_2,\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2\right)-(\mathbf{y}-\hat{\mathbf{y}}_1,\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2)=0\Rightarrow\hat{\mathbf{y}}_1-\hat{\mathbf{y}}_2=0$

命题1（最小二乘）：记$\hat{\mathbf{y}} = P(\mathbf{y} \mid V)$为$\mathbf{y}$在子空间$V \subset R^n$上的投影，则$\| \mathbf{y} - \hat{\mathbf{y}} \|^2 = \min_{\mathbf{u} \in V} \| \mathbf{y} - \mathbf{u} \|^2$

证：对任何$\mathbf{u} \in V$，因为$\hat{\mathbf{y}} - \mathbf{u} \in V$，故$\mathbf{y} - \hat{\mathbf{y}} \perp \hat{\mathbf{y}} - \mathbf{u}$.





所以$\| \mathbf{y} - \mathbf{u} \|^2 = \| \mathbf{y} - \hat{\mathbf{y}} + \hat{\mathbf{y}} - \mathbf{u} \|^2 = \| \mathbf{y} - \hat{\mathbf{y}} \|^2 + \| \hat{\mathbf{y}} - \mathbf{u} \|^2 \geq \| \mathbf{y} - \hat{\mathbf{y}} \|^2$.

投影矩阵

容易知道投影$P(\cdot|V)$是$R^n \to R^n$的线性变换：

$$P(\mathbf{x}a+\mathbf{y}b|V)=P(\mathbf{x}|V)a+P(\mathbf{y}|V)b，\ \forall \mathbf{x},\mathbf{y}\in R^n,\forall \text{实数}a,b$$

故存在$n\times n$矩阵P_V,使得$P(\mathbf{y}|V)=P_V\mathbf{y}$,　P_V称为投影矩阵.

命题2. 设X是$n\times p$列满秩矩阵，则$L(X)\subset R^n$的投影矩阵

$$P_X = X(X^\top X)^{-1}X^\top$$

另外，$I_n-P_X=I_n-X(X^\top X)^{-1}X^\top$是$L(X)$的正交补空间$L(X)^\perp$对应的投影阵

$X_{n\times p}=(\mathbf{x}_1,\ldots,\mathbf{x}_p)$的列向量张成的子空间

$$L(X)=\left\{X\mathbf{b}\,|\,\mathbf{b}\in R^p\right\}=\left\{\mathbf{x}_1b_1+\ldots+\mathbf{x}_pb_p\,|\,b_1,\ldots b_p\in R\right\}$$

证明：对任意$\mathbf{y}$，设它在$L(X)$上的投影为$\hat{\mathbf{y}} = P_X\mathbf{y}$，下面求$P_X$：

(1) 首先，$\hat{\mathbf{y}} \in L(X)$，即存在$\mathbf{b}_{p\times 1}$使得$\hat{\mathbf{y}} = X\mathbf{b}$

(2) 其次，$\mathbf{y} - \hat{\mathbf{y}} \perp L(X)$，特别地垂直于$X$的每一列，即

$$X^\top(\mathbf{y} - \hat{\mathbf{y}}) = 0$$

即$0 = X^\top(\mathbf{y} - X\mathbf{b}) = X^\top\mathbf{y} - X^\top X\mathbf{b} \Rightarrow \mathbf{b} = (X^\top X)^{-1}X^\top\mathbf{y}$,

所以$\hat{\mathbf{y}} = X\mathbf{b} = X\left\{(X^\top X)^{-1}X^\top\mathbf{y}\right\} = \left\{X(X^\top X)^{-1}X^\top\right\}\mathbf{y}$

$$\Rightarrow P_X = X(X^\top X)^{-1}X^\top$$

几个注解

注1：如果$X_{n\times p}$列正交且单位模长，$X^\top X = I_p$，则X的各列$\mathbf{x}_1,\ldots,\mathbf{x}_p$构成$V=L(X)$的标准正交基，向$V=L(X)$投影意味着只需向$X$的各列投影：

$$P_X\mathbf{y} = X(X^\top X)^{-1}X^\top\mathbf{y} = XX^\top\mathbf{y} = \sum_{k=1}^{p}\mathbf{x}_k\left(\mathbf{x}_k^\top\mathbf{y}\right)$$

其中$\mathbf{x}_k^\top\mathbf{y}$为投影坐标，$\mathbf{x}_k$为基。

注2：投影矩阵P_X由$L(X)$唯一决定.

注3：当X不是列满秩的($X^T X$不可逆)，P_{X}一般以广义逆表达，写成$P_{\mathrm{X}} = X(X^\top X)^- X^\top$的形式，其中$A^-$是$A$的任一广义逆，满足$AA^-A=A$.
(虽然广义逆$(X^\top X)^-$不唯一，但$P_{\mathrm{X}} = X(X^\top X)^- X^\top$是唯一的）。

注4：下面如果不做特别说明，我们假设X是列满秩的，但所有后续结论对于不满秩的情形也成立。

投影矩阵左乘一个矩阵

设V是R^n的任一线性子空间，对应的投影矩阵为P_V,任一向量$\mathbf{y}$在V上的投影为$P_V\mathbf{y}$。那么对任一$n\times m$矩阵A, $P_V A$对A的各列做投影：

- $\hat{A}=P_V A$: A各列在V上的投影组成的矩阵

  若A的各列为$A=(\mathbf{a}_1,\ldots,\mathbf{a}_m)$, $P_V A=(P_V\mathbf{a}_1,\ldots,P_V\mathbf{a}_m)$
- $A^{\perp}=A-P_V A=(I-P_V)A$: A各列关于V正交化。

投影矩阵的性质

性质1. (投影阵 $\Leftrightarrow$ 对称幂等阵)

(1) 设 $X_{n\times p}$列满秩, 则投影矩阵$P_X = X(X^\top X)^{-1}X^\top$是对称幂等矩阵.

(2) 反之，若$A_{n\times n}$是任一对称幂等矩阵,则存在$X_{n\times r}$，$r = rank(A)$, 使得 $A = X(X^\top X)^{-1}X^\top$.

证:(1)对于任何$\mathbf{y} \in R^n, \mathbf{z} = P_X\mathbf{y} \in L(X)$，所以$P_X\mathbf{z} = \mathbf{z}$,此即
$P_X^2\mathbf{y} = P_X(P_X\mathbf{y}) = P_X\mathbf{y}$，故$P_X^2 = P_X$.

$$P_X^2 = P_X \Leftrightarrow P_X(I_n - P_X) = 0, (I_n - P_X)^2 = I_n - P_X$$

引理：若A为对称幂等矩阵 $(A=A^{\mathsf{T}}, A^2=A)$，则

（1）其特征根为0或1;

（2）存在正交矩阵O使得$A=O\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}O^{\mathsf{T}}$，其中$r=Rank(A)$,

(3) $\mathrm{tr}(A)=\mathrm{Rank}(A)$

(2). 若A为一个$n\times n$对称幂等阵，$r=rank(A)$,由引理,存在n阶正交阵 $O=(O_1,O_2), O_1$为$n\times r, O_2$为$n\times(n-r)$,使得

$$A=O\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}O^{\mathsf{T}}=(O_1,O_2)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} O_1{}^{\mathsf{T}} \\ O_2{}^{\mathsf{T}} \end{pmatrix}=O_1O_1{}^{\mathsf{T}}$$

但$I_n=O^{\mathsf{T}}O=\begin{pmatrix} O_1{}^{\mathsf{T}} \\ O_2{}^{\mathsf{T}} \end{pmatrix}(O_1,O_2)=\begin{pmatrix} O_1{}^{\mathsf{T}}O_1 & O_1{}^{\mathsf{T}}O_2 \\ O_2{}^{\mathsf{T}}O_1 & O_2{}^{\mathsf{T}}O_2 \end{pmatrix}\Rightarrow O_1{}^{\mathsf{T}}O_1=I_r$

所以$A=O_1O_1{}^{\mathsf{T}}=O_1(O_1{}^{\mathsf{T}}O_1)^{-1}O_1{}^{\mathsf{T}}$是一个投影矩阵。

$A^{\top}B=0$:

A的列与B的列正交

性质2：按列划分$X=(X_1,X_2)$,若$X_1{}^{\top}X_2=0$,则 $P_X=P_{X_1}+P_{X_2}$.

特别地，若 $X_{n\times p}=(\mathbf{x}_1,\ldots,\mathbf{x}_p)$各列正交，则$P_X=P_{\mathbf{x}_1}+\ldots+P_{\mathbf{x}_p}$.

证：$P_X=(X_1,X_2)\begin{pmatrix} X_1{}^{\top}X_1 & \mathbf{0} \\ \mathbf{0} & X_2{}^{\top}X_2 \end{pmatrix}^{-1}\begin{pmatrix} X_1{}^{\top} \\ X_2{}^{\top} \end{pmatrix}=P_{X_1}+P_{X_2}$.

性质3. X为$n\times p$矩阵，若$A_{p\times p}$可逆,则$L(XA)=L(X),P_{XA}=P_X$

证明：$L(XA)\subset L(X)=L(XAA^{-1})\subset L(XA)$,所以$L(XA)=L(X)$,

因而$P_{XA}=P_X$ (直接验证：$P_{XA}=XA[(XA)^{\top}XA]^{-1}(XA)^{\top}=X[X^{\top}X]^{-1}X^{\top}$)

向量的数乘:
向量在左,标量在右

$\mathbf{x}$向量,c标量$c=\mathbf{a}^{\top}\mathbf{b}$

$c\mathbf{x}=\left(\mathbf{a}^{\top}\mathbf{b}\right)\mathbf{x}=\mathbf{a}^{\top}\mathbf{b}\mathbf{x}$ (wrong!)

$\mathbf{x}c=\mathbf{x}\left(\mathbf{a}^{\top}\mathbf{b}\right)=\mathbf{x}\mathbf{a}^{\top}\mathbf{b}$ (right!)

$X_{n\times p}:R^{p}\to R^{n}$,

$X^{T}:R^{n}\to R^{p}$

投影$P=X(X^{\top}X)^{-1}X^{\top}:R^{n}\to R^{n}$

$\mathbf{y}\in R^{n}\xrightarrow{X^{\top}}X^{\top}\mathbf{y}\in R^{p}\xrightarrow{(X^{\top}X)^{-1}}\hat{\boldsymbol{\beta}}\in R^{p}\xrightarrow{X}\hat{\mathbf{y}}\in R^{n}$

$X_1^{\top}X_2=0:X_1$与X_2的列正交

$P_{(X_1,X_2)}=P_{X_1}+P_{X_2}$

A可逆,则$L(XA)=L(X)$

$P_{XA}=P_{X}$

$BC:C$各列的B-变换

$\hat{A}=P_{V}A$: A各列在V上的投影

$A^{\perp}=A-P_{V}A$: $A^{\perp}$各列与V正交。

性质4.(矩阵正交化)：$X=(X_1,X_2),X_1,X_2$分别为$n\times p,n\times q$矩阵，令$X_2^{\perp}=X_2-P_{X_1}X_2$,则$X_2^{\perp}$与$X_1$的各列正交,即$X_1^{\top}X_2^{\perp}=0$，且

$$P_X=P_{(X_1,X_2)}=P_{(X_1,X_2^{\perp})}=P_{X_1}+P_{X_2^{\perp}}$$

证：$(X_1,\ X_2^{\perp})=(X_1,\ X_2)\begin{pmatrix} I_{p_1} & -(X_1^{\top}X_1)^{-1}X_1^{\top}X_2 \\ 0 & I_{p_2} \end{pmatrix}\triangleq(X_1,\ X_2)A$，

其中A可逆，由性质2、3，$P_{(X_1,X_2)}=P_{(X_1,X_2^{\perp})}=P_{X_1}+P_{X_2^{\perp}}$

Gram-Schmidt正交化:

若干$n\times 1$向量$\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3...,$未必相互正交。

Gram-Schmidt正交化:

$$\mathbf{x}_1^{\perp}=\mathbf{x}_1$$

$$\mathbf{x}_2^{\perp}=\mathbf{x}_2-P_{\mathbf{x}_1}\mathbf{x}_2=\mathbf{x}_2-\mathbf{x}_1(\mathbf{x}_1{}^{\mathsf{T}}\mathbf{x}_1)^{-1}\mathbf{x}_1{}^{\mathsf{T}}\mathbf{x}_2$$

$$\mathbf{x}_3^{\perp}=\mathbf{x}_3-P_{(\mathbf{x}_1,\mathbf{x}_2)}\mathbf{x}_3=\mathbf{x}_3-P_{(\mathbf{x}_1^{\perp},\mathbf{x}_2^{\perp})}\mathbf{x}_3=\mathbf{x}_3-P_{\mathbf{x}_1^{\perp}}\mathbf{x}_3-P_{\mathbf{x}_2^{\perp}}\mathbf{x}_3$$

$$=\mathbf{x}_3-\mathbf{x}_1(\mathbf{x}_1{}^{\mathsf{T}}\mathbf{x}_1)^{-1}\mathbf{x}_1{}^{\mathsf{T}}\mathbf{x}_3-\mathbf{x}_2^{\perp}(\mathbf{x}_2^{\perp}{}^{\mathsf{T}}\mathbf{x}_2^{\perp})^{-1}\mathbf{x}_2^{\perp}{}^{\mathsf{T}}\mathbf{x}_3$$

...

卡方分布

定义：$\mathbf{x} \sim N_n(0, I_n)$, 则 $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x} \sim \chi_n^2$

χ_n^2 的pdf：$f(x) = \dfrac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, x > 0,$

命题3：若 $\mathbf{x} \sim N_n(\boldsymbol{\mu}, \Sigma)$, 则 $(\mathbf{x}-\boldsymbol{\mu})^{\mathsf{T}}\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}) \sim \chi_n^2$

证明：$\mathbf{y} = \Sigma^{-1/2}(\mathbf{x}-\boldsymbol{\mu}) \sim N_n(0, \mathrm{I})$,

则 $\mathbf{y}^{\mathsf{T}}\mathbf{y} = (\mathbf{x}-\boldsymbol{\mu})^{\mathsf{T}}\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}) \sim \chi_n^2$

命题4：$\mathbf{x} \sim N_n(0, I_n)$，设$A_{n\times n}$是秩为$r$的对称幂等阵(投影矩阵)，则$\mathbf{x}^\top A\mathbf{x} = \| A\mathbf{x} \|^2 \sim \chi_r^2$

证明：存在正交矩阵O使得$A = O\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} O^\top$

令$\mathbf{y} = O^\top \mathbf{x}$，则$\mathbf{y} \sim N(0, I_n)$。所以

$$\mathbf{x}^\top A\mathbf{x} = \mathbf{x}^\top O\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} O^\top \mathbf{x} = \mathbf{y}^\top \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \mathbf{y} = y_1^2 + \ldots + y_r^2 \sim \chi_r^2$$

注：显然$I_n - A$是秩为$n-r$的对称幂等矩阵，$\| (I_n - A)\mathbf{x} \|^2 \sim \chi_{n-r}^2$

命题5：$\mathbf{x} \sim N_n(0, I_n)$，假设$A_{k\times n}, B_{m\times n}$为常数矩阵，如果$AB^\top = 0$（即$A$各行向量与$B$的各行向量正交），则$A\mathbf{x}$与$B\mathbf{x}$独立。

证明：假设$A_{k\times n}, B_{m\times n}$为行满秩矩阵($AB^\top = 0$, 必定有$k + m \le n$),

$$\begin{pmatrix} A \\ B \end{pmatrix} \mathbf{x} \sim N\left(0, \begin{pmatrix} A \\ B \end{pmatrix} (A^\top, B^\top)\right) \text{ 即 } N\left(0, \begin{pmatrix} AA^\top & 0 \\ 0 & BB^\top \end{pmatrix}\right)$$

所以$A\mathbf{x}$与$B\mathbf{x}$都服从正态分布而且独立

即使A, B不是行满秩，$A\mathbf{x}, B\mathbf{x}$服从退化正态,联合服从退化正态：

$$\begin{pmatrix} A \\ B \end{pmatrix} \mathbf{x} \sim N\left(0, \begin{pmatrix} AA^\top & 0 \\ 0 & BB^\top \end{pmatrix}\right),$$ 此时$A\mathbf{x}, B\mathbf{x}$依旧独立（下面详细证明）

严格的证明：设$A_{k\times n}$的秩为$r\le k, AA^{\top}$为$k\times k$对称阵,由对称矩阵的谱分解定理，

存在标准正交矩阵$Q_{k\times k}=\begin{pmatrix} Q_{1_{r\times k}} \\ Q_2 \end{pmatrix}$使得 $QAA^{\top}Q^{\top}=\begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} Q_1AA^{\top}Q_1^{\top} & Q_1AA^{\top}Q_2^{\top} \\ Q_2AA^{\top}Q_1^{\top} & Q_2AA^{\top}Q_2^{\top} \end{pmatrix}$,

其中Λ_r为r阶对角阵,对角元为$AA^{\top}$的非0特征根。

$$\Rightarrow Q_1AA^{\top}Q_1^{\top}=\Lambda_r,\ Q_2AA^{\top}Q_2^{\top}=0\Rightarrow Q_2A=0\Rightarrow QA=\begin{pmatrix} Q_1A \\ 0 \end{pmatrix}$$

所以$QA\mathbf{x}=\begin{pmatrix} Q_1A\mathbf{x} \\ 0 \end{pmatrix}$，其中$Q_1A\mathbf{x}\sim N(0,Q_1AA^{\top}Q_1^{\top})=N(0,\Lambda_r)$.

同样存在$m\times m$标准正交阵O，使得$OB\mathbf{x}=\begin{pmatrix} O_1B\mathbf{x} \\ 0 \end{pmatrix}$，$O_1B\mathbf{x}\sim N(0,*)$

其中Q_1A, O_1B都是行满秩阵。$AB^{\top}=0\Rightarrow QAB^{\top}O^{\top}=0$，特别地$Q_1AB^{\top}O_1^{\top}=0$

所以$Q_1A\mathbf{x}$与$O_1B\mathbf{x}$相互独立$\Rightarrow QA\mathbf{x}$与$OB\mathbf{x}$相互独立

推论：假设 $\mathbf{x} \sim N_n(0, I_{\mathrm{n}})$,

(1) $A_{n\times n}$对称幂等, B为$m \times n$矩阵，若$BA=0$, 则$B\mathbf{x}$与$\mathbf{x}^{\top}A\mathbf{x}$独立。

(2) $A_{n\times n}, B_{n\times n}$对称幂等，若$BA=0$，则$\mathbf{x}^{\top}A\mathbf{x}$与$\mathbf{x}^{\top}B\mathbf{x}$独立

证明1: (1)由命题5, $B\mathbf{x}$与$A\mathbf{x}$独立，进而与$(A\mathbf{x})^{\top}A\mathbf{x} = \mathbf{x}^{\top}A\mathbf{x}$独立。

（2）由命题5，$A\mathbf{x}$与$B\mathbf{x}$独立，故$\mathbf{x}^{\top}A\mathbf{x}$与$\mathbf{x}^{\top}B\mathbf{x}$独立.

注："对称幂等"条件都可减弱为"对称"

(参见王松桂等：线性模型引论，4.pdf)

(1)$A_{n\times n}$对称, B为$m\times n$矩阵，若$BA=0$, 则$B\mathbf{x}$与$\mathbf{x}^{\mathsf T}A\mathbf{x}$独立。

证明: 存在正交矩阵Q，使得$A=Q\begin{pmatrix}\Lambda_r & 0\\ 0 & 0\end{pmatrix}Q^{\mathsf T}$, 则$\mathbf{x}^{\mathsf T}A\mathbf{x}=\mathbf{x}^{\mathsf T}Q\begin{pmatrix}\Lambda_r & 0\\ 0 & 0\end{pmatrix}Q^{\mathsf T}\mathbf{x}$

$=\mathbf{y}^{\mathsf T}\begin{pmatrix}\Lambda_r & 0\\ 0 & 0\end{pmatrix}\mathbf{y}=\mathbf{y}_1{}^{\mathsf T}\Lambda_r\mathbf{y}_1$, 其中$\mathbf{y}=\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}=Q^{\mathsf T}\mathbf{x}\sim N(0,I_n)$,

由$\mathbf{y}=Q^{\mathsf T}\mathbf{x}\Rightarrow\mathbf{x}=Q\mathbf{y}\Rightarrow B\mathbf{x}=BQ\mathbf{y}\mathrel{\hat=}C\mathbf{y}$.下面说明$B\mathbf{x}$只与$\mathbf{y}_2$有关，即$C$的前$r$行为0.

$$由于0=BA=CQ^{\mathsf T}Q\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}Q^{\mathsf T}=C\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}Q\Rightarrow C\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}=0$$

$$记C=(C_1,C_2),所以(C_1,C_2)\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}=C_1\Lambda\Rightarrow C_1=0$$

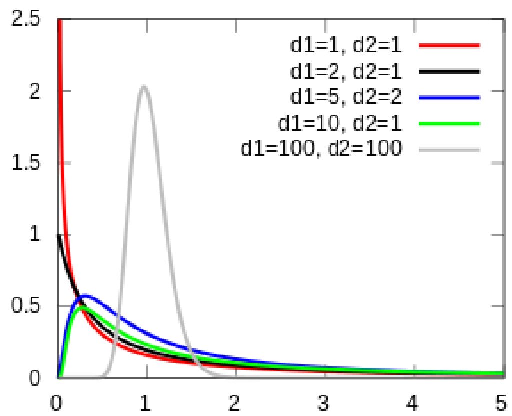
所以，$B\mathbf{x}=C\mathbf{y}=(C_1,C_2)\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}=C_2\mathbf{y}_2$

由于$\mathbf{y}_1$与$\mathbf{y}_2$独立，所以$\mathbf{x}^{\mathsf T}A\mathbf{x}=\mathbf{y}_1{}^{\mathsf T}\Lambda_r\mathbf{y}_1$与$B\mathbf{x}=C_2\mathbf{y}_2$独立。

F 分布

F-分布也叫做Snedecor's F-分布或者Fisher-Snedecor分布，其定义为两个独立平均卡方随机变量之比的分布:

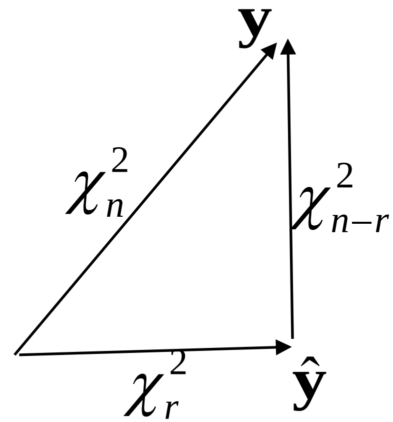
$$F = \frac{\chi_m^2 / m}{\chi_n^2 / n} \sim F_{m,n}$$



命题6：假设$\mathbf{y}\sim N_n(0,I_n)$, V是R^n的r维数子空间，

$\hat{\mathbf{y}}=P_V\mathbf{y}, \mathbf{y}^{\perp}=(I_n-P_V)\mathbf{y}$，则

$$\frac{\|\hat{\mathbf{y}}\|^2/r}{\|\mathbf{y}^{\perp}\|^2/(n-r)}=\frac{\mathbf{y}^{\mathsf{T}}P_V\mathbf{y}/r}{\mathbf{y}^{\mathsf{T}}(I_n-P_V)\mathbf{y}/(n-r)}\sim F_{r,n-r}$$



证明：由命题4，$\|\hat{\mathbf{y}}\|^2\sim\chi_r^2$，$\|\mathbf{y}^{\perp}\|^2\sim\chi_{n-r}^2$，

因为$P_V(I_n-P_V)=0$，由命题5(推论),

$\|\hat{\mathbf{y}}\|^2$ 与 $\|\mathbf{y}^{\perp}\|^2$ 独立,所以命题成立。

- $\|\mathbf{y}\|^2=\|\hat{\mathbf{y}}\|^2+\|\mathbf{y}^{\perp}\|^2\sim\chi_n^2$;
- $\|\hat{\mathbf{y}}\|^2\sim\chi_r^2, \|\mathbf{y}^{\perp}\|^2\sim\chi_{n-r}^2$,两者独立;
- $\dfrac{\|\hat{\mathbf{y}}\|^2/r}{\|\mathbf{y}^{\perp}\|^2/(n-r)}\sim F_{r,n-r}$;
- $\dfrac{\|\hat{\mathbf{y}}\|^2}{\|\mathbf{y}^{\perp}\|^2+\|\mathbf{y}^{\perp}\|^2}\sim Beta\left(\dfrac{r}{2},\dfrac{n-r}{2}\right)$

例1. 假设$x_1,\ldots,x_n$ iid $\sim N(0,\sigma^2)$, $\bar{x}$为样本均值，s^2为样本方差.

令$\mathbf{x}=(x_1,\ldots,x_{\mathrm{n}})^{\mathrm{T}}\sim N(\mathbf{0},\sigma^2 I_n)$, $V=L(\mathbf{1})$

$$\hat{\mathbf{x}}=P_{\mathbf{1}}\mathbf{x}=\mathbf{1}(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T\mathbf{x}=\mathbf{1}\bar{x},$$

$$\mathbf{x}^{\perp}=\mathbf{x}-\hat{\mathbf{x}}=\mathbf{x}-\mathbf{1}\bar{x},\quad \|\mathbf{x}^{\perp}\|^2=\sum(x_i-\bar{x})^2=(n-1)s^2$$

由命题6, $\dfrac{\|\hat{\mathbf{x}}\|^2/r}{\|\mathbf{x}^{\perp}\|^2/(n-r)}=\dfrac{\|\mathbf{1}\bar{x}\|^2}{\|\mathbf{x}^{\perp}\|^2/(n-1)}=\dfrac{n\bar{x}^2}{s^2}\sim F_{1,n-1}.$

注: 这与我们已知的结果$\dfrac{\sqrt{n}\bar{x}}{s^2}\sim t_{n-1}$是一致的，因为$\dfrac{n\bar{x}^2}{s^2}=\left(\dfrac{\sqrt{n}\bar{x}}{s^2}\right)^2$,

而$\left(t_{n-1}\right)^2=\left(\dfrac{N(0,1)}{\sqrt{\chi^2_{n-1}/(n-1)}}\right)^2=\dfrac{\chi^2_1}{\chi^2_{n-1}/(n-1)}=F_{1,n-1}$