4.5 HW 5

作业 5 链接

练习 4.1 若 $\mathbf{x}_{m\times1}, \mathbf{y}_{n\times1}$ 为任意两个随机向量, $E(\mathbf{y}\mid\mathbf{x})$ 称为回归函数, 记 $\boldsymbol{\epsilon}=\mathbf{y}-E(\mathbf{y}\mid\mathbf{x})$.

(a) 证明 $\boldsymbol{\epsilon}$ 与 $\mathbf{x}$ 不相关 (因此, $\mathbf{y}$ 可分解为两个不相关的部分: $\mathbf{y}=E(\mathbf{y}\mid\mathbf{x})+\boldsymbol{\epsilon}$).

(b) 利用 (a) 的结果证明 $\min_{f:R^m\to R^n} E\|\mathbf{y}-f(\mathbf{x})\|^2 = E\|\mathbf{y}-E(\mathbf{y}\mid\mathbf{x})\|^2$ (提示: 首先证明 $\hat{f}(\mathbf{x})=E(\mathbf{y}\mid\mathbf{x})$ 极小化给定 $\mathbf{x}$ 条件下的条件期望 $E\left(\|\mathbf{y}-f(\mathbf{x})\|^2\mid\mathbf{x}\right)$.

(c) 利用 (a) 的结果证明方差矩阵分解公式: $\operatorname{var}(\mathbf{y})=\operatorname{var}[E(\mathbf{y}\mid\mathbf{x})]+E[\operatorname{var}(\mathbf{y}\mid\mathbf{x})]$ (提示: 只需证明 $\operatorname{var}(\boldsymbol{\epsilon})=E[\operatorname{var}(\mathbf{y}\mid\mathbf{x})]$).

(d) 假设 $\mathbf{x},\mathbf{y},\mathbf{z}$ 是三个随机向量, 证明

$$\operatorname{cov}(\mathbf{x},\mathbf{y})=\operatorname{cov}[E(\mathbf{x}\mid\mathbf{z}),E(\mathbf{y}\mid\mathbf{z})]+E[\operatorname{cov}(\mathbf{x},\mathbf{y}\mid\mathbf{z})].$$

证明

(a)

$$\begin{aligned}
&\operatorname{Cov}(\varepsilon,x)\\
=&\operatorname{Cov}(y-E(y\mid x),x)\\
=&E\left((y-E(y\mid x))x^\top\right)-E(y-E(y\mid x))Ex^\top\\
=&E\left((y-E(y\mid x))x^\top\right)\\
=&0\quad\text{（根据条件期望定义）}
\end{aligned}$$

(b)

$$\begin{aligned}
&E\left(\|y-f(x)\|^2\mid x\right)\\
=&E\left(\|y-E(y\mid x)+E(y\mid x)-f(x)\|^2\mid x\right)
\end{aligned}$$

$$\begin{aligned}
&\text{记}\varepsilon=y-E(y\mid x)\\
&\text{由于}\varepsilon\text{与}x\text{ 不相关}\\
\Rightarrow&\operatorname{Cov}(\varepsilon,E(y\mid x)-f(x))=0\\
\Rightarrow&E\left(\varepsilon(E(y\mid x)-f(x))^\top\right)=0
\end{aligned}$$

$$\begin{aligned}
&E\left(\|y-f(x)\|^2\mid x\right)\\
=&E\left(\|\varepsilon\|^2\mid x\right)+E\left(\|E(y\mid x)-f(x)\|^2\mid x\right)\\
=&E\left(\|\varepsilon\|^2\mid x\right)+\|E(y\mid x)-f(x)\|^2\\
\geqslant&E\left(\|\varepsilon\|^2\mid x\right)
\end{aligned}$$

$$\begin{aligned}
&E\|y-f(x)\|^2\\
=&E\left(E\left(\|y-f(x)\|^2\mid x\right)\right)\\
\geqslant&E\left(E\left(\|\varepsilon\|^2\mid x\right)\right)\\
=&E\|\varepsilon\|^2\\
&\text{当且仅当}f(x)=E(y\mid x)
\end{aligned}$$

(c)

$$
\begin{aligned}
&\operatorname{Var}(\varepsilon)\\
=&E\left(\varepsilon\varepsilon^{\top}\right)-(E\varepsilon)(E\varepsilon)^{\top}\\
=&E\left(E\left(\varepsilon\varepsilon^{\top}\mid x\right)\right)\\
=&\ E((y-E|y|x))(y-E(y\mid x))^{\top}\mid x)\\
=&E\left(E\left(yy^{\top}+E(y\mid x)E(y\mid x)^{\top}-2yE(y\mid x)^{\top}\mid x\right)\right)\\
=&E\left(E\left(yy^{\top}\mid x\right)+E(y\mid x)E(y\mid x)^{\top}-2E(y\mid x)E(y\mid x)^{\top}\right)\\
=&\ E\left(E\left(yy^{\top}\mid x\right)-E(y\mid x)E(y\mid x)^{\top}\right)\\
=&E(\operatorname{Var}(y\mid x))
\end{aligned}
$$

(d)

$$
\begin{aligned}
&\operatorname{Cov}(E(x\mid z),E(y\mid z))\\
=&E\left(E(x\mid z)E(y\mid z)^{\top}\right)-ExEy^{\top}
\end{aligned}
$$

$$
\begin{aligned}
&E(\operatorname{Cov}(x,y\mid z))\\
=&E\left(E(xy^{\top}\mid z)-E(x\mid z)E(y\mid z)^{\top}\right)\\
=&Exy^{\top}-E\left(E(x\mid z)E(y\mid z)^{\top}\right)
\end{aligned}
$$

$$
\begin{aligned}
&\operatorname{Cov}(E(x\mid z),E(y(z))+E(\operatorname{Cov}(x,y\mid z))\\
=&Exy^{\top}-ExEy^{\top}\\
=&\operatorname{Cov}(x,y)
\end{aligned}
$$

练习 4.2 假设 y 是一元响应变量, $\mathbf{x}$ 是 q 元随机向量, z 是 $(p-q)\times 1$ 随机向量, 都是可观测变量。我们感兴趣的问题是研究 y 与 $\mathbf{x}$ 的关系, 而 $\mathbf{z}$ 是干扰因素 (既与 y 有关也与 $\mathbf{x}$ 有关), 需要加以控制。假设如下线性模型

$$y=a+\mathbf{b}^{\top}\mathbf{x}+\mathbf{c}^{\top}\mathbf{z}+\epsilon,\epsilon\sim\left(0,\sigma^{2}\right),\epsilon\perp\mathbf{x},\mathbf{z}(1)$$

问题是, 该模型是否正确地实现了目标 “在控制 $\mathbf{z}$ 的条件下, 研究 y 与 $\mathbf{x}$ 的关系” ? 为了考察模型 (1) 中加项 $\mathbf{c}^{\top}\mathbf{z}$ 的作用, 以及考察系数 $\mathbf{b}$ 的含义, 我们令

$$y^{\perp}=y-\Sigma_{y\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z},\quad \mathbf{x}^{\perp}=\mathbf{x}-\Sigma_{\mathbf{xz}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z},$$

则可以证明模型 (1) 蕴含了

$$y^{\perp}=a+\mathbf{b}^{\top}\mathbf{x}^{\perp}+\epsilon,\quad \epsilon\sim\left(0,\sigma^{2}\right),\quad \epsilon\perp\mathbf{x}^{\perp}.(2)$$

这说明模型 (1) 中线性项 $\mathbf{c}^{\top}\mathbf{z}$ 确实起到了控制（或消除）$\mathbf{z}$ 影响的作用。进一步, 可以证明

$$\mathbf{b}=\Sigma_{\mathbf{xx\bullet z}}^{-1}\Sigma_{\mathbf{x}y\bullet\mathbf{z}},(3)$$

特别地, 当 $q=1$ 时, $b\propto\rho_{xy\bullet z}$ 。结果 (4) 说明了模型 (2) 中 $\mathbf{x}$ 的系数 $\mathbf{b}$ 确实代表了偏相关系数, 干扰因素的影响通过在模型 (1) 添加 $\mathbf{c}^{\top}\mathbf{z}$ 得到了消除。

另一方面, 假设模型确实是正确模型, 但假如我们没有意识到 z 是一个潜在的干扰因素, 并假设模型

$$y=\alpha+\beta^{\top}\mathbf{x}+\delta,\delta\sim\left(0,\tau^{2}\right),\delta\perp\mathbf{x}(4)$$

也就是说, 我们将模型 (1) 中的 $\mathbf{c}^{\top}\mathbf{z}$ 一项归入了误差项:

$$\delta=\epsilon+\mathbf{c}^{\top}\mathbf{z}-E\left(\mathbf{c}^{\top}\mathbf{z}\right)$$

其中最后一项 $-E\left(\mathbf{c}^{\top}\mathbf{z}\right)$ 是为了保证 $E(\delta)=0$, 而 $\alpha=a+E\left(\mathbf{c}^{\top}\mathbf{z}\right),\beta=\mathbf{b}$ 。显然 δ 与 $\mathbf{x}$ 不独立, 那么基于

错误模型 (4), 由课件命题 2

$$\mathbf{b} = \beta = \Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}, (5)$$

与 (3) 相比, 这个表达没有控制 $\mathbf{z}$, 因而是错误的表达。本题的任务是, 假设模型 (1) 是正确模型, 证明 (2) 和 (3)。

证明

(2)

$$\begin{aligned} y^{\perp} &= y - \Sigma_{yz}\Sigma_{zz}^{-1}z \\ &= a + b^{\top}x + c^{\top}z + \varepsilon - \Sigma_{yz}\Sigma_{zz}^{-1}z \\ &= a + b^{\top}\left(x^{\perp} + \Sigma_{xz}\Sigma_{zz}^{-1}z\right) + c^{\top}z + \varepsilon - \Sigma_{yz}\Sigma_{zz}^{-1}z \\ &= a + b^{\top}x^{\perp} + \varepsilon + c'^{\top}z \end{aligned}$$

$$\begin{aligned} &\text{由于} y^{\perp} x^{\perp}, \varepsilon \text{与} z \text{ 不相关} \\ \Rightarrow & c' = 0, \quad \varepsilon \perp x^{\perp} \\ \Rightarrow & y^{\perp} = a + b^{\top}x^{\perp} + \varepsilon \end{aligned}$$

(3)

$$\varepsilon = y^{\perp} - a - b^{\top}x^{\perp} \text{与} x^{\perp} \text{ 不相关}$$

$$\operatorname{Cov}\left(\varepsilon, x^{\perp}\right) = 0$$

$$\begin{aligned} &\Sigma_{yx\cdot z} - b^{\top}\Sigma_{xx\cdot z} = 0 \\ &b = \Sigma_{xx\cdot z}^{-1}\Sigma_{xy\cdot z} \end{aligned}$$