Skitse til ShSt opgave 3.14

(a) Hvis vi har givet x (altså betinget af x) så kan vi let se hvilken værdi af $g(x)$, som jo så er konstant, der minimerer den betingede middelværdi **givet** x:

$$\text{MSE}(x) \equiv \text{E}\left[(y-g(x))^2|x\right].$$

Thi ovenstående betingede middelværdi er givet ved

$$\begin{aligned}\text{MSE}(x) &= \text{E}(y^2|x) + \text{E}(g(x)^2|x) - 2\text{E}(g(x)y|x) \\ &= g(x)^2 - 2\text{E}(y|x)g(x) + \text{E}(y^2|x).\end{aligned}$$

Bemærk, at når x er givet så er $g(x)$ også givet (konstant). Man skal så indse, at $\text{MSE}(x)$ er mindst for $g(x) = \text{E}(y|x)$. Det er let, for udtrykket ovenfor er et andengrads polynomium i $g(x)$ (med grenene opad), som vi ved fra elementær matematik antager sit minimum, når $g(x) = \frac{-(-2\text{E}(y|x))}{2\cdot 1} = \text{E}(y|x)$.

Altså har vi for ethvert x, at

$$\text{MSE}(x) \geq \text{E}\left[(y-\text{E}(y|x))^2|x\right]. \tag{1}$$

Vi skal nu bruge "Law of Iterated Expectatons" (LIE), der siger, at middelværdien af den betingede middelværdi er lig med den ubetingede middelværdi. Anvender man nemlige LIE på x og $(y-g(x))^2$ får vi altså

$$\text{MSE} = \text{E}\left(y-g(x)\right)^2 = \text{E}\left(\text{E}(y-g(x))^2|x\right) = \text{E}\left(\text{MSE}(x)\right).$$

(Dette resultat er nøjagtig det samme som *Hint*'et i opgaven).

Fra (1) får vi så,

$$\begin{aligned}\text{MSE} = \text{E}\left(\text{MSE}(x)\right) &\geq \text{E}\left(\text{E}\left[(y-\text{E}(y|x))^2|x\right]\right) \\ &= \text{E}\left[y-\text{E}(y|x)\right]^2.\end{aligned}$$

Altså kan vi se, at MSE er mindst ved at vælge $g(x) = \text{E}(y|x)$.

(b) Hvis $y = x^2 + z$, og x og z er uafhængige Gaussiske, med $\text{E}x = \text{E}z = 0$ og $\text{V}x = \text{V}z = 1$, så er $\text{E}(y|x) = x^2 + \text{E}(z|x) = x^2$. Derfor følger fra spørgsmål **(a)**, at

$$\text{MSE} = \text{E}\left((y-x^2)^2\right) = \text{E}(z^2) = 1.$$

(c) Her er $g(x)$ restringeret til at være på formen $g(x) = a + bx$. At finde funktionen $g(x)$ der minimerer MSE svarer så til at finde værdierne af a og b som minimerer MSE. Først bemærkes det at $\text{E}(y) = 1$ fra (b). Vi får så

$$\begin{aligned}\text{minimize MSE} &= \text{E}\left(y-a-bx\right)^2 \\ &= \text{E}(y^2) + a^2 + b^2\text{E}(x^2) + 2ab\text{E}(x) - 2a\text{E}(y) - 2b\text{E}(xy) \\ &= \text{E}(y^2) + a^2 + b^2\text{E}(x^2) + 2ab\text{E}(x) - 2a - 2b\text{E}(xy).\end{aligned}$$

Denne er mindst når de afledede mht. hhv. a og b er nul, det vil sige

$$\begin{aligned} 2a + 2b\mathrm{E}(x) - 2 &= 0 \\ 2\mathrm{E}(x^2)b + 2a\mathrm{E}(x) - 2\mathrm{E}(xy) &= 0, \end{aligned}$$

hvilket giver

$$\begin{aligned} a &= 1 \\ b &= \frac{\mathrm{E}(xy)}{\mathrm{E}(x^2)} = \frac{\mathrm{E}(x^3)}{\mathrm{E}(x^2)} = 0. \end{aligned}$$

$MSE = \mathrm{E}[(y-1)^2] = \mathrm{E}[y^2] - 2\mathrm{E}[y] + 1 = \mathrm{E}[x^4] + \mathrm{E}[z^2] - 2 + 1 = 3 + 1 - 2 + 1 = 3.$

Ekstra:

I tidsrækkeanalyse er $y = x_{t+1}$ og $x = x_t$ og overstående er en ekstra grund til vi forecaster x_t vha. $\mathrm{E}(x_{t+1}|x_t)$: Udover at give mening rent intuitivt, så minimerer det også den kvadrerede fejl i middel.