

# Oppgaver, STK4040, uke 42

17. oktober 2005

## Oppgave 1

Oppgave 6.2.1 i boka. (Tips: P er lik M fra forrige ukes oppgave 1, så vi vet f.eks. allerede at den er symmetrisk og idempotent.)

## Oppgave 2

Verifiser ligning (6.5.6) i boka, og vis at definisjonen i ligning (6.5.7) er ekvivalent.

## Oppgave 3

Anta  $\varepsilon \sim N(0, 1)$  uavhengig av  $\mathbf{x} \sim N_5(\boldsymbol{\mu}, \Sigma)$ , der

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ -2 \end{pmatrix} \quad \text{og} \quad \Sigma = \begin{pmatrix} 1.0 & 0.5 & 0.5 & -0.2 & 0.0 \\ 0.5 & 1.5 & 0.8 & 0.1 & 1.0 \\ 0.5 & 0.8 & 1.0 & -0.7 & 0.5 \\ -0.2 & 0.1 & -0.7 & 2.0 & -0.5 \\ 0.0 & 1.0 & 0.5 & -0.5 & 2.0 \end{pmatrix},$$

La  $y = 3 + \mathbf{x}^t \boldsymbol{\beta} + \varepsilon$ , der og  $\boldsymbol{\beta} = (1, -1, 2, 0.5, 0.3)^t$ .

Trekk 20 observasjoner  $(\mathbf{X}, \mathbf{y})$  fra  $(\mathbf{x}^t, y)$  og gjør en regresjon av  $\mathbf{y}$  på  $\mathbf{X}$ . Beregn standardavvik og kovariansmatrise til de estimerte regresjonskoeffisientene, og sammenlign med de sanne verdiene.

Sentrer så  $\mathbf{X}$ , og gjenta regresjonen. Sammenlign resultatene med resultatene for usentrerte data. Hva er forskjellig, og hvorfor? Er det best å sentrere eller ikke?