

# Oppgaver, STK4040, uke 39

26. september 2005

## Oppgave 1

Bruk datasettet fra sist uke, og gjør en «manuell» prinsipalkomponentanalyse (PCA) på det; dvs. via egenverdidekomposisjon av kovariansmatrisen.

Sammenlign resultatene (varianser, skårer og ladninger) med resultatene fra den «innebygde» PCA-funksjonen (R: `prcomp()`). Er det noen forskjeller, og isåfall hva skyldes de?

Blir det noen forskjell om man bruker egenverdidekomposisjonen av  $S_u$  eller  $S$ ? Isåfall, hva består forskjellen i, og hvilken versjon er «riktig»?

## Oppgave 2

Bruk datasettet på filen `datasett_39.txt`. Gjør PCA på dataene, og tolk som vanlig. Standardiser datasettet slik at variablene har varians 1, og gjør PCA på det standardiserte datasettet. Hva blir tolkningen nå? Hvorfor blir det forskjell? Hva gir «riktig» bilde? Når skal vi standardisere?