

Oppgaver - SOK3023

Maskinl ring for  konomer

Laget av Markus J. Aase

January 13, 2025

Dette dokumentet inneholder oppgaver studentene i SOK-3023 kan jobbe med etter f rste uka av kurset. En del av svarene vil dere kunne finne fra de fysiske forelesningene og/eller kompendiet. Mens noen sp rsm l krever at dere opps ker informasjonen selv i dokumentasjon til Tensorflow eller andre sted p  internett.

1. Gitt f lgende matriser

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

- (a) Utf r matrisemultiplikasjonen $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$. Husk at elementet C_{ij} beregnes ved   ta skalarproduktet av rad i i \mathbf{A} og kolonne j i \mathbf{B} .
- (b) Forklar hvorfor dimensjonene til resultatmatrisen \mathbf{C} blir 2×2 .
- (c) Verifiser beregningene ved   finne hvert element i \mathbf{C} :

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}.$$

2. Oppgave: Fra tabell til matrise og vektor

Tabellen nedenfor viser informasjon om fire personer, inkludert deres alder, antall  r med utdanning, og om de jobber fulltid (ja = 1, nei = 0). L nnen er oppgitt i tusen kroner.

Person	Alder	Utdanning (�r)	Fulltid (1/0)	L�nn (Y, tusen kr)
1	25	16	1	450
2	30	14	0	350
3	40	18	1	600
4	35	12	0	400

Svar p  f lgende oppgaver:

- (a) Skriv de uavhengige variablene som en matrise \mathbf{X} , der hver rad representerer en person, og hver kolonne representerer en av variablene.
- (b) Skriv målvariabelen som en vektor \mathbf{Y} .
- (c) Verifiser at dimensjonene til \mathbf{X} og \mathbf{Y} er henholdsvis 4×3 og 4×1 .

Hint:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

3. **Tensorflow** I dette faget kommer vi til å bruke *Tensorflow*, som er et populært rammeverk for maskinlæring som støtter bruk av både CPU og GPU for å akselerere beregninger.

Svar på følgende oppgaver:

- (a) Hva er en rank-0, rank-1, rank-2 og rank-3 tensor? Gjerne illustrer. Hva vil en rank-4 tensor være?
- (b) Tensorflow er optimalisert for å gjøre beregninger med matriser og tensorer. Hva er parallellbehandling? Hva er fordelene?
- (c) Forklar forskjellen mellom hvordan TensorFlow bruker CPU og GPU til beregninger. Hvorfor kan GPU være mer effektiv for dype nevralt nettverk sammenlignet med CPU?
- (d) Anta at du har installert TensorFlow på en maskin med både CPU og GPU. Hvordan kan du sjekke hvilke enheter TensorFlow bruker for å utføre beregningene? Gi et eksempel på en TensorFlow-kommando som lister tilgjengelige enheter.

4. **Normer i lineær algebra:** Gitt en vektor $\mathbf{v} = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$:

- (a) Beregn ℓ_1 -normen $\|\mathbf{v}\|_1 = \sum_{i=1}^n |v_i|$.
- (b) Beregn ℓ_2 -normen $\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$.
- (c) Sammenlign de to normene og diskuter hvordan ℓ_1 - og ℓ_2 -normer vektlegger elementene i vektoren.

5. **Mean Squared Error (MSE) som en norm:** Gitt dataene (\hat{y}_i, y_i) som representerer prediksjoner og sanne verdier for i -te datapunkt:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2.$$

- (a) Forklar hvordan MSE kan relateres til ℓ_2 -normen. Hint: Se på vektoren $\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$,

der $r_i = \hat{y}_i - y_i$, som residualvektoren.

Definisjon: Residualvektor

La \mathbf{y} være en vektor som representerer de faktiske verdiene (observasjoner) og $\hat{\mathbf{y}}$ være en vektor med modellens prediksjoner. Da defineres residualvektoren \mathbf{r} som:

$$\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}},$$

hvor elementene i \mathbf{r} er gitt ved:

$$r_i = y_i - \hat{y}_i, \quad \text{for } i = 1, 2, \dots, n.$$

Tolkning:

- $r_i > 0$: Modellen undervurderte verdien (prediksjonen er lavere enn observasjonen).
- $r_i < 0$: Modellen overvurderte verdien (prediksjonen er høyere enn observasjonen).
- $r_i = 0$: Modellen predikerte nøyaktig riktig for dette datapunktet.

- (b) Vis at MSE kan uttrykkes som:

$$\text{MSE} = \frac{1}{n} \|\mathbf{r}\|_2^2.$$

6. **Sammenligning mellom ℓ_p -normer og MSE:** Gitt residualvektoren $\mathbf{r} = \begin{bmatrix} -1 \\ 2 \\ -3 \\ 4 \end{bmatrix}$:

- (a) Beregn ℓ_1 -normen $\|\mathbf{r}\|_1$, ℓ_2 -normen $\|\mathbf{r}\|_2$, og ℓ_∞ -normen $\|\mathbf{r}\|_\infty = \max |r_i|$.
 (b) Diskuter hvordan de forskjellige normene gir ulik informasjon om residualene.
 (c) Beregn MSE for residualvektoren, og forklar hvordan den er relatert til ℓ_2 -normen.

7. **Praktisk anvendelse:** Du jobber med en regresjonsmodell som gir følgende prediksjoner:

$$\hat{\mathbf{y}} = \begin{bmatrix} 2.1 \\ 3.9 \\ 6.0 \\ 8.2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix}.$$

- (a) Beregn residualvektoren $\mathbf{r} = \hat{\mathbf{y}} - \mathbf{y}$.
- (b) Finn $\|\mathbf{r}\|_1$, $\|\mathbf{r}\|_2$, og MSE.
- (c) Diskuter hvordan en lav $\|\mathbf{r}\|_1$ eller $\|\mathbf{r}\|_2$ kan være et mål for modellens kvalitet.