# Øvingsforelesning til 12 - Interpolasjon, regresjon og markovkjeder

## Interpolasjon og regresjon

**1.** Gitt et ligningssystem $A\mathbf{x} = \mathbf{b}$ som ikke har en løsning. Jeg kan bruke minste kvadratersmetode for å

- minimere lengden til vektoren $\mathbf{x}$
- finne den beste approksimasjon til en løsning for systemet
- jeg vet ikke fordi jeg ikke har forberedt meg.

**2. (Vår 2020, oppgave 31)** Finn linjen som går gjennom eller passer best til punktene i $\mathbb{R}^2$

$$\begin{bmatrix}-1\\-1\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}.$$

**3.** La $A$ være en $4 \times 2$-matrise. Hvilken av matrisene kan være lik $A^T \cdot A$?

$$L = \begin{bmatrix}14 & 6\\6 & 4\end{bmatrix}, \ M = \begin{bmatrix}14 & 6\\4 & 4\end{bmatrix}, \ N = \begin{bmatrix}14 & 0\\6 & 14\end{bmatrix}$$

**4.** Vi skal finne polynomer som går gjennom eller passer best til punktene i $\mathbb{R}^2$

$$\begin{bmatrix}0\\-1\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix}, \begin{bmatrix}3\\20\end{bmatrix}.$$

**a)** Finnes det en linje på formen $y = ax + b$ som går gjennom alle punktene? Hvis ikke, bruk minste kvadraters metode til å finne linjen som passer best til punktene.

**b)** Det finnes ikke et annengradspolynom som går gjennom alle punktene heller. Hvilket ligningssystem må vi se på for å vise dette? Hva må vi sjekke for dette systemet?

**c)** Men det finnes et unikt tredjegradspolynom som går gjennom alle punktene. Sett opp et ligningssystem for koeffisientene til dette polynomet, og finn koeffisientene.

**d)** Kan du nå gjette hva $x$ må være for at den følgende setningen er sant: Gitt $m$ datapunkter i $\mathbb{R}^2$. Da finnes det et unikt $x$-tegradsplynom som går gjennom alle punktene. Hva ville du gjort for å vise at $x$en du valgte er den riktige?

## Markovkjeder

**1.** Markovkjeder er

- en viktig type stokastiske prosesser
- ikke
- jeg vet ikke fordi jeg ikke har forberedt meg.

**2.** En likevektsvektor

- legger like mye vekt på alle sine koordinater
- er det samme som nullvektoren
- beskriver tilstanden en Markovkjede konvergerer til.

**3.** Hvilken av vektorene kan være en likevektsvektor for en Markovkjede?

$$\mathbf{u} = \begin{bmatrix}1/3\\4/3\\-2/3\end{bmatrix}, \ \mathbf{v} = \begin{bmatrix}1/2\\1/2\\1/2\end{bmatrix}, \ \mathbf{w} = \begin{bmatrix}1/3\\1/2\\1/6\end{bmatrix}$$

**4.** Hvilken av matrisene er en stokastisk matrise

$$L = \begin{bmatrix}1 & 1\\2 & 1\end{bmatrix}, \ M = \begin{bmatrix}1/3 & 1\\2/3 & 0\end{bmatrix}, \ N = \begin{bmatrix}1/3 & 4/3\\2/3 & -1/3\end{bmatrix}$$

**5.** Er følgende stokastiske matriser regulære? Begrunn svaret ditt.

**a)** $P = \begin{bmatrix}0.2 & 1\\0.8 & 0\end{bmatrix}$

**b)** $Q = \begin{bmatrix}1 & 0.2\\0 & 0.8\end{bmatrix}$

**6. (Høst 2020, oppgave 6)** For å simulere proteinene som danner taggene til koronaviruset SARS-CoV-2 bruker forskere overgangsalgoritmer (for eksempel i prosjektet Folding@Home ved Stanford). De trenger din hjelp til å teste om algoritmen de har implementert fungerer som den skal, og har derfor forenklet utregningene slik at du kan gjøre dem for hånd.

Proteinet S kan innta tre ulike tilstander, A, B og C. Hvert mikrosekund, altså hvert $10^{-6}$-sekund, kan den bytte tilstand.

- Gitt at S er i tilstand A, og man venter et mikrosekund så er det: 2/3 sannsynlighet for at S fortsatt er i tilstand A, 1/3 sannsynlighet for at S går til tilstand B, 0 sannsynlighet for at S går til tilstand C.
- Gitt at S er i tilstand B, og man venter et mikrosekund så er det: 0 sannsynlighet for at S går til A, 3/4 sannsynlighet for at S fortsatt er i tilstand B, 1/4 sannsynlighet for at S går til tilstand C.
- Gitt at S er i tilstand C, og man venter et mikrosekund så er det: 1/3 sannsynlighet for at S går til A, 1/3 sannsynlighet for at S går til tilstand B, 1/3 sannsynlighet for at S fortsatt er i tilstand C.

**a)** Finn den stokastiske matrisen M som representerer hvordan S endrer tilstander for hvert mikrosekund.

**b)** Gitt at S er i tilstand A, hva er sannsynligheten for at S er i tilstand B to (2) mikrosekunder senere?

**c)** Forskerene simulerer at proteinet S får endre tilstand fritt, slik det vil, i et helt sekund. Estimer sannsynligheten for at S er i hver av de 3 tilstandene på slutten av sekundet. Hvilken tilstand er det mest sannsynlig at S vil være i?