11.6 Markovkjeder

En Markovkjede blir brukt til å beskrive et eksperiment som utføres flere ganger på samme måte og hvor utfallet av hvert eksperiment bare er avhengig av utfallet av forrige forsøk og ikke av tidligere forsøk.

En elementær modell for værvarsling: Finne ut været det skal bli i morgen fra været observert i dag i Lisboa. Modellen bygger på følgende antakelse:

- Hvis det i dag er sol, er det 90% sannsynlighet for sol i morgen og 10% sanns. at det skal bli regn;
- Hvis det i dag er regn, er det 50% sanns. for sol i morgen og 50% sanns. for regn.

Dvs.

sol i dag $\xrightarrow{90\%}$ sol i morgen

sol i dag $\xrightarrow{10\%}$ regn i morgen

regn i dag $\xrightarrow{50\%}$ sol i morgen

regn i dag $\xrightarrow{50\%}$ regn i morgen

Disse opplysningene kan vi sette opp i en såkalt overgangsmatrise.

$$
\begin{array}{ccc}
 & \text{Utfall i dag} \rightarrow \quad \text{sol} \quad \text{regn} & \text{utfall i morgen} \\
P = & \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix} & \begin{array}{c} \text{sol} \\ \text{regn} \end{array}
\end{array}
$$

Def. Generelt: Overgangsmatrisa er en $k \times k$ sannsynlighetsmatrise $P = (p_{ij})_{\substack{i=1\ldots k \\ j=1\ldots k}}$, der k er antall mulige utfall $\{1, \ldots, k\}$ og p_{ij} er sannsynligheten fra utfall j i forrige forsøk at man skal få utfall i.

I eksempelet vårt er sol utfall 1 og regn utfall 2 og f. eks. $p_{21} = 0.1$ er sannsynligheten for at hvis det er sol i dag, regner det i morgen.

* Merk: $p_{11} + p_{21} = 1$ og $p_{12} + p_{22} = 1$

Betyr at hvis vi har sol i dag, så kommer vi enten til å få sol (p_{11}) eller regn (p_{21}) i morgen, s.a. $p_{11} + p_{21}$ dekker alle mulige utfall.

- Samme resonnement gjelder for alle søyler i en overgangsmatrise, dvs. $p_{1j} + p_{2j} + \ldots + p_{kj} = 1$

- Vi kan bruke P til å varsle været i dager framover ut ifra været vi har i dag.
- Anta at situasjonen i dag (dvs. været observert i dag) er beskrevet av en vektor.

$$x^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{array}{l} \text{sol} \\ \text{regn} \end{array}$$

Da er situasjonen i morgen beskrevet av vektor

$$x^{(1)} = Px^{(0)} = \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix} \begin{array}{l} \text{sol} \\ \text{regn} \end{array}$$

Dette stemmer med vår antakelse, dvs. i morgen er det sol med 90% sanns. og regn med 10% sanns.

- Hva med dagen etter?

$$x^{(2)} = Px^{(1)} = \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix} \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 0.86 \\ 0.14 \end{pmatrix}$$
$$= P^2 x^{(0)}$$

$\Rightarrow$ sol med 86% sanns.

Generelt:
$$x^{(n)} = P x^{(n-1)}$$
$$x^{(n)} = P^n x^{(0)}$$

Merk: $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ er sannsynlighetsvektorer, dvs.

$$x_1^{(n)} + x_2^{(n)} + \dots + x_k^{(n)} = 1$$

Hvis prosessen konvergerer når $n \to \infty$ har vi en likevektsvektor. $p = \lim_{n\to\infty} x^{(n)}$.

Prosessen konvergerer ikke alltid, men når P er regulær da kalles også Markovkjeden regulær og vi har en likevektsvektor.

Def. En overgangsmatrise er regulær hvis det fins $e \in \mathbb{N}$ s.a. $P^e = \underbrace{P \cdot P \cdots P}_{n \text{ ganger}}$ har bare positive (>0) elementer.

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ er ikke regulær. $I^n = I \quad \forall n$.

Teorem 11.6.3

P er en regulær overgangsmatrise og x en sannsynlighetsvektor, da

$$P^n x \xrightarrow[n\to\infty]{} q,$$

der q er en sannsynlighetsvektor, uavhengig av n, med alle komponenter positive. (uten bevis).

Anta nå at $x^{(n)} = Px^{(n-1)}$ er en Markovkjede og q er en likevektsvektor. Da siden

$$x^{(n)} \to q$$
$$Px^{(n-1)} \to Pq$$

$$
\begin{array}{ccc}
x^{(n)} & = & Px^{(n-1)} \\
\downarrow & & \downarrow \\
q & = & Pq
\end{array}
$$

$$q - Pq = 0$$
$$(I - P)q = 0$$

og man kan også bevise at q er en unik sannsynlighetsvektor, s.a. $Pq = q$ (Teorem 11.6.4).

I eksempelet om været har vi

$$(I - P)q = 0 \qquad \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix} \right) q = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0.1 & -0.5 \\ -0.1 & 0.5 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 1 & -5 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\Rightarrow q_1 - 5q_2 = 0 \qquad q_2 = s, \; q_1 = 5s \;\Rightarrow\; \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 5s \\ s \end{pmatrix}$ og siden q er en sannsynlighetsvektor er $q_1 + q_2 = 1 \Rightarrow 5s + s = 1 \Rightarrow s = 1/6$

$\Rightarrow \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0.8333 \\ 0.1667 \end{pmatrix}$. Dette betyr at i det lange løp er 83% av dagene med sol og 17% med regn.

Eksempel 2: Meningsmåling om EU.

Anta at blant norges befolkning gir en meningsmåling følgende situasjon per i dag:

$$x^{(0)} = \begin{pmatrix} 0.50 \\ 0.30 \\ 0.20 \end{pmatrix} \begin{array}{l} \text{ja} \\ \text{nei} \\ \text{vet ikke} \end{array}$$

Og trenden over de siste måneder gir følgende overgangsmatrise:

$$P = \begin{array}{c} \begin{array}{ccc} \text{ja} & \text{nei} & \text{vet ikke} \end{array} \\ \begin{pmatrix} 0.85 & 0.05 & 0.05 \\ 0.01 & 0.85 & 0.2 \\ 0.14 & 0.1 & 0.75 \end{pmatrix} \end{array} \begin{array}{l} \text{ja} \\ \text{nei} \\ \text{vet ikke} \end{array}$$

Anta at opinionsutviklingen er slik at overgangsmatrisa holder fra måned til måned. Hva er situasjonen om 1, og 2 måneder? Og i det lange løp?

$$x^{(1)} = Px^{(0)} = \begin{pmatrix} 0.85 & 0.05 & 0.05 \\ 0.01 & 0.85 & 0.2 \\ 0.14 & 0.1 & 0.75 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.3 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 0.45 \\ 0.3 \\ 0.25 \end{pmatrix}$$

$$x^{(2)} = Px^{(1)} = \begin{pmatrix} \quad & \quad & \quad \end{pmatrix} \begin{pmatrix} 0.45 \\ 0.3 \\ 0.25 \end{pmatrix} = \begin{pmatrix} 0.4100 \\ 0.3095 \\ 0.2805 \end{pmatrix}$$

I det lange løp har vi q: $Pq = q$, dvs. $(I-P)q = 0$

$$(I-P)q = \begin{pmatrix} 0.15 & -0.05 & -0.05 \\ -0.01 & 0.15 & -0.2 \\ -0.14 & -0.1 & 0.25 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \quad \begin{pmatrix} 15 & -5 & -5 \\ -1 & 15 & -20 \\ -14 & -10 & 25 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \quad \begin{matrix} 3q_1 = q_2 + q_3 \\ \\ -q_1 + 15q_2 - 20q_3 = 0 \end{matrix}$$

$q_1 = \frac{35}{44}s$

$q_2 = \frac{61}{44}s$

$q_3 = s$

q er en sannsynlighetsvektor:

$$q_1 + q_2 + q_3 = 1$$

$$\frac{35 + 61 + 44}{44} \cdot s = 1 \Leftrightarrow s = \frac{44}{140} = \frac{11}{35}$$

$$q = \begin{pmatrix} 0.25 \\ 0.43 \\ 0.31 \end{pmatrix}$$

THE END...