

## KONVERGENS AV MARKOVKJEDER (SEC. 5.9)

Vi har indentifisert to viktige spørsmål i teorien av Markovkjeder:

- (1) Finnes der alltid en likevektsvektor?
- (2) Hvis den finnes, hvornår konvergerer tilstandsvektorer  $\vec{x}_k$  mod likevektsvektoren?

Vi kommer til at svare på begge disse spørsmål.

**Eksistens av likevektsvektorer.** Vi kan starte med et nemt teorem:

**Teorem 10.** *Hvis  $P$  er en stokastisk matrise så er 1 et egenverdi for  $P$ .*

*Bevis.* Husk at  $P$  og  $P^T$  har de samme egenverdier fordi deres karakteristiske polynomier er de samme: For hvert  $\lambda$  er

$$\det(\lambda I_n - P^T) = \det(\lambda I_n - P).$$

La  $\vec{e}$  være vektoren med 1 i hvert indgang, dvs.

$$\vec{e} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

og huske at hvert søyle i den stokastiske matrise  $P = (\vec{s}_1, \dots, \vec{s}_n)$  summer opp til 1, dvs. søylevektorer  $\vec{s}_1, \dots, \vec{s}_n$  er sannsynlighetsvektorer. Nu kan vi se at

$$P^T \vec{e} = \begin{pmatrix} -\vec{s}_1^T - \\ \vdots \\ -\vec{s}_n^T - \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n (\vec{s}_1)_i \\ \vdots \\ \sum_{i=1}^n (\vec{s}_n)_i \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = 1 \cdot \vec{e}.$$

Det viser at 1 er en egenverdi for  $P^T$  og vi konkluderer at den er også en egenverdi for  $P$ .  $\square$

Teoremet er ikke helt nok for at vise at der finnes en likevektsvektor for hvert Markovkjede. Det er ikke klart at der finnes en egenvektor til egenverdien 1 som er også en sannsynlighetsvektor (den skal ha positive indganger som oppsummerer til 1).

**Teorem.** *Hvert stokastisk matrise har en likevektsvektor.*

*Bevis.* Teorem 10 viser at der finnes en egenvektor  $\vec{x}$  til egenverdien 1 for  $P$ . Anta at

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

med tall  $x_1, \dots, x_n$ , og definer vektoren

$$|\vec{x}| = \begin{pmatrix} |x_1| \\ \vdots \\ |x_n| \end{pmatrix}$$

med positive indganger. Vi kommer til at vise at

$$(1) \quad P|\vec{x}| = |\vec{x}|,$$

dvs. at  $|\vec{x}|$  er også en egenvektor til egenverdien 1 for  $P$ . Husk at  $|\vec{x}|$  kan ikke være nullvektoren fordi  $\vec{x}$  er ikke nullvektoren (som egenvektor) og fordi  $|x_i| = 0$  hvis og kun hvis  $x_i = 0$ . Når vi har vist (1) kan vi definere

$$\vec{q} = \frac{1}{\sum_{k=1}^n |x_k|} |\vec{x}|,$$

som er en likevektsvektor for  $P$  og vi ville være færdig.

For at vise (1) definerer vi

$$\vec{y} = P|\vec{x}| - |\vec{x}|,$$

og (1) holder hvis og kun hvis  $\vec{y} = \vec{0}$ . For hvert index  $i$  har vi

$$|x_i| = |(P\vec{x})_i|,$$

fordi  $P\vec{x} = \vec{x}$ . Nu kan vi bruke trekantuligheden og at  $P$  har positive indganger for at vise

$$|x_i| = |(P\vec{x})_i| = \left| \sum_{j=1}^n P_{ij}x_j \right| \leq \sum_{j=1}^n P_{ij}|x_j| = (P|\vec{x}|)_i$$

Denne ulighed viser at

$$y_i = (P|\vec{x}|)_i - |x_i| \geq 0.$$

Alle indganger av  $\vec{y}$  er positiv! Men fordi stokastiske matriser bevarer summer av vektorer har vi

$$\sum_{i=1}^n (P|\vec{x}|)_i = \sum_{i=1}^n |x_i|,$$

og vi finner at

$$\sum_{i=1}^n y_i = \sum_{i=1}^n (P|\vec{x}|)_i - \sum_{i=1}^n |x_i| = 0.$$

Fordi en summe av positive tall (som  $y_i$ ) er kun like 0 hvis alle tallene er 0 konkluderer vi at  $\vec{y} = \vec{0}$ , og det viser (1).  $\square$

Alle stokastiske matriser har en likevektsvektor. Men er den også unik? Den næste eksempel viser at likevektsvektorer i generelle tilfald er *ikke* unik!

**Eksempel.** La oss komme tilbake til migrasjonseksemplen. Men nu vil vi modellere migrasjon mellom Oslo, resten av Norge og Tyskland. Det betyr at vi har tillstander

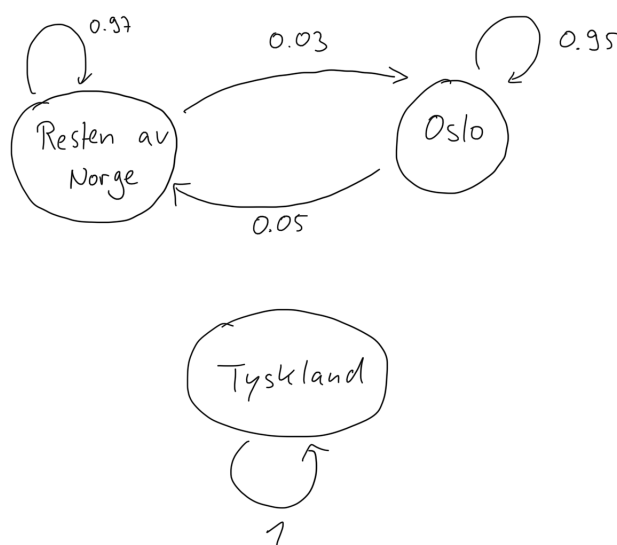
$$\mathcal{S} = \{\text{resten av Norge, Oslo, Tyskland}\}.$$

La oss anta at tysker er meget mærkelig og de vil ikke flytte til Norge eller andre steder hen. De vil bar blive i Tyskland hele tiden. Fordi tysker er så merklig vil Nordmenn heller ikke flytte til tyskland og vi har sannsynligheder som i Figure 1.

Den tilsvarende stokastiske matrise er

$$P = \begin{pmatrix} 0.95 & 0.03 & 0 \\ 0.05 & 0.97 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

som inneholder den stokastiske matrise fra sidste eksempel som blok. Har Markovkjeden som er definert med  $P$  et likevektsvektor? Selvfølgelig har det det. Hvis der



FIGUR 1. Fiktivt migrasjon mellom Norge og Tyskland.

ikke finnes noen som bor i tyskland så kan vi bare bruke den gamle likevektsvektor og finne en likevektsvektor

$$\vec{q} = \begin{pmatrix} 0.375 \\ 0.625 \\ 0 \end{pmatrix}.$$

Men vi kan også anta at der bor noen mennesker i tyskland. For eksempel kunne halvdelen av mennesker i modellen bor i tyskland, og den norske populasjon kunne være i likevekt mellom Oslo og ikke Oslo. Det ville svar til likevektsvektoren

$$\vec{q} = \frac{1}{2} \begin{pmatrix} 0.375 \\ 0.625 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.1875 \\ 0.3125 \\ 0.5 \end{pmatrix}$$

Med sammen prinsip kan man konstruere mange likevektsvektorer. Hvis vi regner ut egenverdier og egenvektorer for den stokastiske matrise  $P$  så finner vi at  $\vec{q}$  og  $\vec{q}$  utspenner den 2-dimensjonale egenrom til egenverdien 1.

Hvornår finnes der en unike likevektsvektor? I den neste avsnit får vi et svar.

**Konvergens for regulære stokastiske matriser.** La  $P$  være en  $n \times n$  stokastiske matrise, og  $\vec{x}_0, \vec{x}_1, \vec{x}_2, \dots$  stokastiske vektorer slik at

$$\vec{x}_{k+1} = P \vec{x}_k,$$

for alle  $k \in \mathbb{N}$ . Alternativ, vi kan skrive

$$\vec{x}_k = P^k \vec{x}_0.$$

Vi har set i migrasjonseksemplen (uten tyskland) at der fantes en unik likevektsvektor  $\vec{q}$  slik at

$$P^k \vec{x}_0 \rightarrow \vec{q},$$

hvis  $k \rightarrow \infty$ , for alle sannsynlighetsvektorer  $\vec{x}_0$ . For hvilke stokastiske matriser  $P$  har vi denne slaks konvergens? Hvornår finnes der en sannsynlighetsvektor  $\vec{x}_\infty$  slik at

$$P^k \vec{x}_0 \rightarrow \vec{x}_\infty,$$

hvis  $k \rightarrow \infty$ , for alle sannsynlighedsvektorer  $\vec{x}_0$ ? Hvis det er sant for alle sannsynlighedsvektorer  $\vec{x}_0$ , så er det også sant for en likevektsvektor  $\vec{q}$  for  $P$  som finnes ifølge teoremet fra sidste seksjon. Men hvis vi velger  $\vec{x}_0 = \vec{q}$  så holder

$$P^k \vec{x}_0 = P^k \vec{q} = \vec{q} \rightarrow \vec{x}_\infty,$$

hvis  $k \rightarrow \infty$ , hvis og kun hvis  $\vec{x}_\infty = \vec{q}$ . Vi konkluderer at hvis følger  $P^k \vec{x}_0$  konvergerer imod en vektor  $\vec{x}_\infty$  for alle sannsynlighedsvektorer  $\vec{x}_0$ , så er  $\vec{x}_\infty = \vec{q}$  en unik likevektsvektor for  $P$ . Hvorfor unikt? Hvis der finnes en annen likevektsvektor  $\vec{q}'$  så har vi jo

$$P^k \vec{q}' = \vec{q}' \rightarrow \vec{x}_\infty = \vec{q},$$

hvis  $k \rightarrow \infty$ , og det kan kun være sant hvis  $\vec{q} = \vec{q}'$ . Fra eksemplerne som vi har set før kunne vi gætte at hvert stokastiske matrise med en unikt likevektsvektor har denne konvergenssegenskap. Det er *ikke* rigtigt! La oss se på en eksempel hvor det kan gå gjæld.

**Eksempel** (Trafiklys). La oss se på en Markovkjede som modellerer et trafiklys. Der er tre tilstander

$$\mathcal{S} = \{\text{Grøn}, \text{Gul}, \text{Rød}\},$$

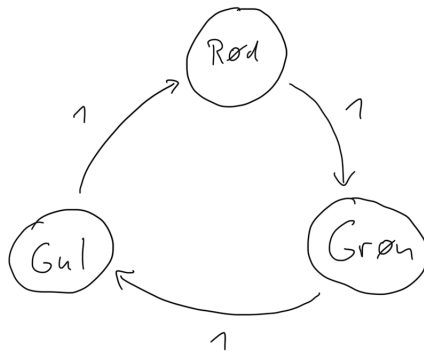
og sannsynligheder

$$p(\text{Grøn} \rightarrow \text{Gul}) = p(\text{Gul} \rightarrow \text{Rød}) = p(\text{Rød} \rightarrow \text{Grøn}) = 1.$$

Denne Markovkjede er faktisk deterministisk, og der er ikke nogen tilfeldighet i denne system (godt i trafikken). Den tilsvarende stokastiske matrise er

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

hvor 'Rød' er den første tilstand, 'Gul' den annen og 'Grøn' den tredje.



FIGUR 2. Et trafiklys som Markovkjede.

Hva er likevektsvektoren i denne eksempel? Det er nemt at se at

$$\vec{q} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix},$$

er den eneste likevektsvektor i denne eksempel<sup>1</sup>. Men i denne eksempel konvergerer heller ikke alle følger  $\vec{x}_0, \vec{x}_1, \dots$  med  $\vec{x}_{k+1} = P\vec{x}_k$  for stokastiske vektorer  $\vec{x}_0$ . For eksemplene, hvis

$$\vec{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

så får vi

$$\vec{x}_1 = P\vec{x}_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

$$\vec{x}_2 = P\vec{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$\vec{x}_3 = P\vec{x}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\vec{x}_4 = P\vec{x}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

⋮

og Markovkjeden går bar igennem de samme tre vektorer igen og igjen (lige som et trafiklys). Selv om der finnes kun et likevektsvektoren konvergerer Markovkjeden ikke for alle startvektorer!

For hvilke stokastiske matriser konvergerer følgerne  $P^k\vec{x}_0$  imod en unik likevektsvektor  $\vec{q}$  for alle sannsynlighetsvektorer  $\vec{x}_0$ ? En mulighet for at vise denne slaks konvergens er at huske potensmetoden. Hvis den stokastiske matrise  $P$  har 1 som dominante egenverdi slik at alle andre egenverdier har absolutverdi strengt mindre en 1, så konvergerer  $P^k\vec{x}_0$  imod en egenvektor til egenverdien 1 som er den unik likevektsvektor  $\vec{q}$ . Matrisen i trafiklyseksempel har faktisk egenverdier  $\lambda_1 = 1, \lambda_2 = \cos(\pi/3) + i\sin(\pi/3)$  og  $\lambda_3 = \overline{\lambda_2} = \cos(\pi/3) - i\sin(\pi/3)$  som alle har absolutverdi 1. Derfor er  $\lambda_1 = 1$  ikke en dominant egenverdi i denne eksempel og potensmetoden virker ikke.

Det er ikke så nemt (men mulig) at analysere egenverdier for stokastiske matriser og at vise konvergens mod likevektsvektor med potensmetoden. I følgende diskusjon velger vi en tilgang som er lidt nemmer og hvor vi finner en egenskap for stokastiske matriser som er tillstrækkelig for denne slaks konvergens. La oss starte med at definere denne egenskap:

**Definisjon.** En stokastiske matrise  $P$  kalles regulær hvis der finnes en  $k \in \mathbb{N}$  slik at matrisen  $P^k$  har kun strengt positive indganger.

Vi kan se på noen eksempler som av stokastiske matriser som er regulær og andre som er ikke:

### Eksempel.

- (1) Stokastiske matrisen

$$P = \begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix},$$

<sup>1</sup>Fordi hvis indgangerne ville ikke være de samme så kan den ikke bli det samme under multiplikasjon med  $P$  som syklisk permuterer indgangerne.

fra migrasjonseksempel uten tyskland er regulær fordi den kun har strengt positive indganger (i definisjonen kan vi velge  $k = 1$ ).

- (2) Den stokastiske matrise

$$P = \begin{pmatrix} 0.97 & 0.05 & 0.1 \\ 0 & 0.9 & 0.05 \\ 0.03 & 0.05 & 0.85 \end{pmatrix}$$

er også regulær. Matrisen  $P$  selv har en indgang som er 0, men vi kan regne ut at

$$P^2 = \begin{pmatrix} 0.9439 & 0.0015 & 0.0546 \\ 0.0985 & 0.8125 & 0.089 \\ 0.1845 & 0.0875 & 0.728 \end{pmatrix}$$

som kun har strengt positive indganger. Vi kan velge  $k = 2$  i definisjonen, og det viser at matrisen  $P$  er regulær.

- (3) Den stokastiske matrise

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

fra trafiklyseksempel er *ikke* regulær. Vi kan regne ut at

$$P^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$P^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P^4 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = P,$$

og syklen starter igjen. Der finnes ingen  $k$  slik at  $P^k$  har kun strengt positive indganger.

- (4) Den stokastiske matrise

$$P = \begin{pmatrix} 0.95 & 0.03 & 0 \\ 0.05 & 0.97 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

i migrasjonseksemplene med tyskland er heller *ikke* regulær! Hvis vi regner ut de første potenser

$$P^2 = \begin{pmatrix} 0.904 & 0.0576 & 0 \\ 0.096 & 0.9424 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P^3 = \begin{pmatrix} 0.86168 & 0.082992 & 0 \\ 0.13832 & 0.917008 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

⋮

og vi ser at nullerne i den tredje søyle og tredje rad aldri kommer til at forsvinne.

Hva betyr det konseptuelt hvis en stokastisk matrise er regulær? Det er ikke svært å vise (prøve at se det selv, eller se på eksempler) at indgangen  $(P^k)_{ij}$  er sannsynligheten for at gå fra tilstand  $i$  til tilstand  $j$  i precis  $k$  trinn. En stokastisk

matrise som beskriver en Markovkjede er regulært hvis der finnes en  $k$  slik at for hvert par  $(i, j)$  av tilstander man kan komme fra  $i$  til  $j$  i  $k$  trinn med strengt positiv sannsynlighet. Vær opmarksom på den precise formulering her. I trafiklys eksempel finnes der en  $k_{12}$  slik at man kan komme fra tilstand 1 til tilstand 2 i  $k_{12}$  trinn, og der finnes også en  $k_{13}$  slik at man kan komme fra tilstand 1 til tilstand 3 i  $k_{13}$  trinn. Men  $k_{12} \neq k_{13}$  og der finnes ikke en enkelt naturlig tall  $k$  slik at man kan komme fra hvert  $i$  til hvert  $j$  i  $k$  trinn med strengt positiv sannsynlighet.

Nu hvor vi forstår hva regularitet betyr for stokastiske matriser kan vi vise vores hovedresultat:

**Teorem 11.** *La  $P$  være en  $n \times n$  regulær stokastisk matrise.*

- (1)  $P$  har en unik likevektsvektor  $\vec{q}$ .
- (2) For alle inisjaltilstander  $\vec{x}_0$ , så vil Markovkjeden

$$\vec{x}_{k+1} = P \vec{x}_k,$$

konvergere mot  $\vec{q}$ .

For at vise Theorem 11 er det nemmest at bruke en annen distanse mellom vektorer. For en vektor  $\vec{y} \in \mathbb{R}^n$  introduserer<sup>2</sup> vi

$$\|\vec{y}\|_1 = \sum_{i=1}^n |y_i|.$$

Det er nemt at se noen egenskaper av  $\|\cdot\|_1$ :

- Vi har  $\|\vec{y}\|_1 = 0$  hvis og kun hvis  $\vec{y} = \vec{0}$ .
- Hvis  $\|\vec{x}_k - \vec{q}\|_1 \rightarrow 0$  hvis  $k \rightarrow \infty$  så konvergerer  $\vec{x}_k \rightarrow \vec{q}$  hvis  $k \rightarrow \infty$ .

Den første punkt er sand fordi  $|a| = 0$  hvis og kun hvis  $a = 0$ , og en summe av positive tall er null hvis og kun hvis alle tall i summen er null. Den annen punkt er sand fordi  $\|\vec{x}_k - \vec{q}\|_1 \rightarrow 0$  betyr at indganger av  $\vec{x}_k$  konvergerer imod indganger av  $\vec{q}$ .

Vi har også bruk for en annen definisjon:

**Definisjon.** For en  $n \times n$  matrise  $A$  med søyler  $\vec{s}_1, \dots, \vec{s}_n$  definerer vi

$$\text{diam}(A) = \max_{i,j} \|\vec{s}_i - \vec{s}_j\|_1,$$

som vi kaller for diameteren av  $A$ .

Hvorfor er vi interessert i denne merkelige diameter? Motivasjonen kommer fra den etterfølgende lemma som vi ikke viser her (for en bevis se den annen vedlegg til forelesningen).

**Lemma 1.** For hvert  $n \times n$  matrise  $A$  og hvert vektor  $\vec{y}$  med  $\sum_{k=1}^n y_k = 0$  har vi

$$\|A \vec{y}\|_1 \leq \frac{\text{diam}(A)}{2} \|\vec{y}\|_1.$$

Hvis vi kunne vise at  $\text{diam}(P) < 2$  for en regulær stokastiske matrise, så ville lemmaen vise en slaks kontraksjonsegenskap. Vi kan vise det etterfølgende lemma:

**Lemma 2.** For hvert stokastiske matrise  $P$  har vi  $\text{diam}(P) \leq 2$  og vi har  $\text{diam}(P) < 2$  hvis alle indgangerne i  $P$  er strengt positiv.

<sup>2</sup>Det her må man gerne huske for senere. Distansen  $\|\vec{x} - \vec{y}\|_1$  mellom sannsynlighetsvektorer  $\vec{x}$  og  $\vec{y}$  kalles også *totale variasjons distanse* eller bare *1-norm distanse*. Det er en viktig distanse i statistikk og sannsynlighetsteori.

*Bevis.* Hvis  $a, b \geq 0$  så er det klart at  $|a - b| \leq a + b$ . Derfor har vi

$$\|\vec{s}_i - \vec{s}_j\|_1 = \sum_{k=1}^n |P_{ki} - P_{kj}| \leq \sum_{k=1}^n (P_{ki} + P_{kj}) = 1 + 1 = 2,$$

for søylerne  $\vec{s}_i$  og  $\vec{s}_j$  i  $P$ . Vi konkluderer at

$$\text{diam}(P) = \max_{i,j} \|\vec{s}_i - \vec{s}_j\|_1 \leq 2.$$

Hvis  $a, b > 0$  så har vi  $|a - b| < a + b$  og alle uligheter er strengt hvis alle indgangerne i  $P$  er strengt positiv.  $\square$

*Bevis av Markov's theorem.* La  $P$  være en  $n \times n$  stokastiske matrise og la oss anta at alle indganger i  $P$  er strengt positiv (den generelle tilfald vises på næsten det sammen måle  $\rightsquigarrow$  Se annen vedlegg til forelesningen for detaljer). Lemma 2 viser at

$$\kappa := \frac{\text{diam}(P)}{2} < 1.$$

Fordi stokastiske matriser bevarer summen av indganger i vektorer har vi

$$\sum_{i=1}^n (P\vec{y})_i = 0,$$

hvis

$$\sum_{i=1}^n y_i = 0.$$

Derfor kan vi bruke Lemma 1 flere gang i træk og det viser at

$$\|P^l \vec{y}\|_1 \leq \kappa^l \|\vec{y}\|_1,$$

for alle  $\vec{y}$  slik at

$$\sum_{i=1}^n y_i = 0.$$

La oss nu anta at  $\vec{q}$  er en likevektsvektor for  $P$  (der finnes alltid så en vektor ifølge teoremet fra før). For hvert stokastisk vektor  $\vec{x}_k$  i følgen  $\vec{x}_0, \vec{x}_1, \vec{x}_2, \dots$  finner vi at

$$\vec{y}_k = \vec{q} - \vec{x}_k$$

har indgangerne som summerer til 0 fordi indgangerne i  $\vec{q}$  og i  $\vec{x}_k$  summerer til 1. Vi konkluderer at

$$\begin{aligned} \|\vec{q} - \vec{x}_m\|_1 &= \|P(\vec{q} - \vec{x}_{m-1})\|_1 \\ &= \|P\vec{y}_{m-1}\|_1 \leq \kappa \|\vec{y}_{m-1}\|_1 = \kappa \|\vec{q} - \vec{x}_{m-1}\|_1. \end{aligned}$$

Hvis vi itererer det så finner vi at

$$\|\vec{q} - \vec{x}_m\|_1 \leq \kappa \|\vec{q} - \vec{x}_{m-1}\|_1 \leq \kappa^2 \|\vec{q} - \vec{x}_{m-2}\|_1 \leq \dots \leq \kappa^m \|\vec{q} - \vec{x}_0\|_1,$$

og vi konkludere at

$$\|\vec{q} - \vec{x}_m\|_1 \leq \kappa^m \|\vec{q} - \vec{x}_0\|_1 \rightarrow 0,$$

hvis  $m \rightarrow \infty$  (fordi  $\kappa < 1$ ).

Til sidst kan vi vise at likehetsvektoren er unikt. Anta at  $\vec{q}'$  er også en likevektsvektor for  $P$ . Fra første del av beviset ved vi at følgen  $P^m \vec{q}'$  konvergerer mot likevektsvektoren  $\vec{q}$ , men og fordi  $P^m \vec{q}' = \vec{q}'$  for alle  $m$  finner vi at  $\vec{q}' = \vec{q}$ . Det viser at likevektsvektoren er unike.  $\square$