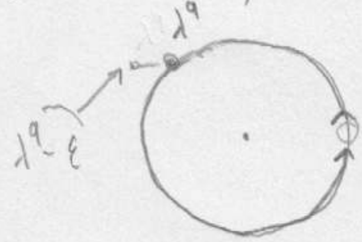
Satz von Perron-Frobenius

Sei $Q \geq 0$ primitiv, ϱ der Spektralradius von Q (Betrag des betragsgrößten Eigenwerts). Dann hat Q genau einen Eigenwert λ mit $|\lambda| = \varrho$ und es gilt $\lambda = \varrho$. Die algebraische (und die geometrische) Vielfachheit von λ ist 1 und "der" Eigenvektor von Q zu λ hat nur positive Einträge.

Bew:

Sei ϱ der Spektralradius von Q. OBdA $\varrho = 1$ (Sonst Übergang zu Q/ϱ.) Sei λ mit $|\lambda| = 1$ ein Eigenwert von Q. Annahme: $\lambda \neq 1$, d.h. $\mathrm{Re}(\lambda) < 1$.

Nach der 4. Beh.: $Q^q > 0 \quad \forall q \geq q^{(0)}$. Da $|\lambda| = 1$, $\lambda \neq 1$ $\exists q$ (oBdA $q \geq q^{(0)}$) mit $\mathrm{Re}(\lambda^q) < 0$.



Wegen $Q^q > 0$ ex. $\varepsilon > 0$ mit $Q^q - \varepsilon I > 0$ (komponentenweise)

Laut Übung ist $\varrho(Q) = \lim\limits_{p \to \infty} \|Q^p\|_\infty^{1/p}$

Wegen $0 < Q^q - \varepsilon I \leq Q^q$ folgt $\|Q^q - \varepsilon I\|_\infty \leq \|Q^q\|_\infty$ (Zeilensummen-norm)

bzw. $(Q^q - \varepsilon I)^p \leq Q^{qp} \Rightarrow \|(Q^q - \varepsilon I)^p\|_\infty \leq \|Q^{qp}\|_\infty$

$$\varrho(Q^q - \varepsilon I) = \lim_{p \to \infty} \|(Q^q - \varepsilon I)^p\|_\infty^{1/p} \leq \lim_{p \to \infty} \|(Q^q)^p\|_\infty^{1/p} = \varrho(Q^q) = 1.$$

Dabei hat $Q^q - \varepsilon I$ den Eigenwert $\lambda^q - \varepsilon$ mit $|\lambda^q - \varepsilon| > 1$ ↯

Also gibt es nur einen Eigenwert vom Betrag 1 und der ist $\lambda = 1$.

$\Rightarrow$ Für allgemein gewählte $v^{(0)} \geq 0$ konvergiert die Potenzmethode,

$v^{(\ell+1)} := Qv^{(\ell)} / \|Qv^{(\ell)}\|_2$ gegen einen Eigenvektor $v \geq 0$

zu $\lambda = 1$. Sei μ eine Komponente mit $v_\mu > 0$.

Wegen $v = \underbrace{Q^q}_{>0} v \geq v_\mu (Q^q)_{:,\mu} > 0$ folgt $v > 0$. ⊛

- Wenn der Eigenraum von Q die Dimension ≥ 2 hätte, so hätte der Nullraum von $Q - I \in \mathbb{R}^{n \times n}$ eine reelle Basis und es gäbe ein $w \in \mathbb{R}^n$ lin. unabh. zu obigem v mit $Qw = w$. Dabei hat w eine Komponente > 0 (sonst $w \to -w$). Sei α maximal, so dass $u := v - \alpha w \geq 0$. $(u \neq 0)$

  $\Rightarrow$ $u \geq 0$ ist ein Eigenvektor von Q zu $\lambda = 1$, aber $u \not> 0$, im Widerspruch zu ⊛.

- Annahme der Jordan-Block zu $\lambda = \rho$ $(=1)$ hat Dimension >1.
- Für $A \geq 0$ und $v > 0$ gilt $\underbrace{\max_i \sum_j A_{ij} v_j}_{= \|Av\|_\infty} \geq \max_i \min_j \{v_j\} \sum_j A_{ij}$

  $= \min_j \{v_j\} \; \|A\|_\infty$

  $\Rightarrow \|A\|_\infty \leq \dfrac{\|Av\|_\infty}{\min_j \{v_j\}}$.

  Für $A = Q^q$ folgt $\|Q^q\|_\infty \leq \dfrac{\|Q^q v\|_\infty}{\min_j \{v_j\}} \overset{Qv = v}{=} \dfrac{\|v\|_\infty}{\min_j \{v_j\}}$ unabh. von q ⊛⊛

  Falls Q einen Jordan-Block zu $\rho = 1$ der Dimension ≥ 2 hätte so wäre $\|JNF^q\|_\infty = \|S Q^q S^{-1}\|_\infty \leq \|S\|_\infty \|Q^q\|_\infty \|S^{-1}\|_\infty$

  $$\| \begin{pmatrix} \ddots & & \\ 1 & 1 & \\ & 1 & \ddots \end{pmatrix}^q \|_\infty \xrightarrow{q \to \infty} \infty \quad \text{(Übung)}$$

  Somit folgt $\|Q^q\|_\infty \xrightarrow{q \to \infty} \infty$ ↯ zu ⊛⊛

Simulated Annealing am Beispiel des TSP-Problems

Sei $\mathfrak{X}$ eine Menge mit $m \gg 1$ Elementen und $f: \mathfrak{X} \to \mathbb{R}$ eine "Bewertungsfunktion".

Bsp: Ein Graph mit Knoten $1, \dots, n$ und gerichteten Kanten (μ, ν) mit Kantenlängen $d_{\mu\nu}$ $(1 \le \mu, \nu \le n,\ \mu \ne \nu,\ n \ge 3)$

Falls die Kante (μ, ν) nicht existiert, so setze

$$d_{\mu\nu} = \sum_{(r,s) \in R} |d_{rs}| + 1.$$

TSP: Gesucht ist eine kürzeste Rundtour in G, die jeden Knoten genau einmal besucht.

$\mathfrak{X}$ = Menge der Rundtouren, $|\mathfrak{X}| = (n-1)! =: m$

(oBdA. fängt die Rundtour in "1" an, suche Nachfolger aus den übrigen Knoten und wiederhole)

$x = (x_1, \dots, x_n)$ sei eine Rundtour, d.h. $x_\rho \in \{1, \dots, n\}$ für $1 \le \rho \le n$ und $x_\mu \ne x_\nu$ für $\mu \ne \nu$.

$$f(x) := \sum_{\mu=1}^{n-1} d_{x_\mu x_{\mu+1}} + d_{x_n x_1} \xrightarrow{\ !\ } \min$$

$\boxed{\rightsquigarrow \text{oBdA } d_{\mu\nu} \ge 0 \ \forall \mu,\nu}$

(Für das TSP gibt es deutlich bessere Verfahren, nachfolgend soll das Prinzip des Simulated Annealing "simuliertes Ausglühen" erklärt werden.)

Sei $\mathfrak{X} = \{x^{(1)}, \dots, x^{(m)}\}$ mit $f(x^{(i)}) \le f(x^{(i+1)})$ für $1 \le i \le m-1$.

(Die Nummerierung ist nicht explizit bekannt, aber sie existiert.)

Ansatz:

- Fixiere eine Matrix $W \in \mathbb{R}^{m \times m}$ mit $W_{ij} \geq 0 \quad \forall i,j$,

  $\sum_{j=1}^{m} W_{ij} = 1$ für $1 \leq i \leq m$.

  (Auch W wird nicht explizit gebildet!)

- Wähle eine "Temperatur" $T > 0$.

- Im Schritt k sei das Element $x^{(i)}$ gegeben.

  Für $1 \leq j \leq m$

  Setze
  $$P_{ij} = \begin{cases} W_{ij} & \text{falls } i > j \text{ d.h. falls } f(x^{(i)}) > f(x^{(j)}) \\ 1 - \sum\limits_{l \neq j}^{m} P_{il} & \text{falls } i = j \\ \underbrace{W_{ij} \cdot \exp\left(\frac{f(x^{(i)}) - f(x^{(j)})}{T}\right)}_{\leq 1} & \text{falls } i < j \end{cases}$$

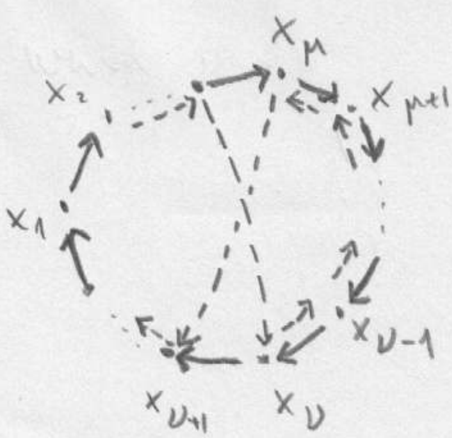
  und wähle das Element $x^{(j)}$ mit Wahrscheinlichkeit P_{ij} als Element in Schritt $k+1$.

  (Und auch P wird nicht explizit gebildet!)

Bsp: Aus $x = x^{(i)}$ werden 2 Indices $\mu \neq \nu$ gleichverteilt gewürfelt. Als "benachbarte" Tour lässt sich dann

$(x_1, \ldots, x_{\mu-1}, x_\nu, x_{\nu-1}, \ldots, x_{\mu+1}, x_\mu, x_{\nu+1}, \ldots, x_n)$ definieren,

falls $\mu < \nu$,



$\longrightarrow x^{(i)}$

$\dashrightarrow x^{(j)}$

und $(x_n, x_{n-1}, \dots, x_{\mu+1}, x_\nu, x_{\nu+1}, \dots, x_\mu, x_{\nu-1}, \dots, x_n)$ falls $\mu > \nu$. (48)

Die Anzahl der benachbarten Touren ist $n(n-1)$

d.h. $w_{ij} = \frac{1}{n(n-1)}$ falls i zu j benachbart ist.

Aufwand: • "Würfeln von μ, ν" $\underline{O(1)}$

• Berechne der Kosten der neuen Tour "x^{neu}" $\underline{O(n)}$

$$\tilde{p}_K := \begin{cases} \exp\left(\frac{f(x) - f(x^{neu})}{T}\right) < 1 & \text{falls } f(x) < f(x^{neu}) \\ 1 & \text{sonst} \end{cases}$$

$\underline{O(1)}$

• Mit WS'keit $\tilde{p}_K$ ersetze x durch x^{neu}

$\underline{O(n)}$

$\downarrow$

(erzeuge Zufallszahl in $[0,1]$ & teste ob diese "$\leq \tilde{p}_K$" ist.)

Analyse:

Sei $v^{(k)} \in \mathbb{R}^m$ der Vektor dessen Komponenten $v_l^{(k)}$ angeben, dass $x^{(l)}$ im Schritt k mit Wahrscheinlichkeit $v_l^{(k)}$ gegeben ist, d.h. $v_l^{(k)} \geq 0$, $\sum_{l=1}^m v_l^{(k)} = 1$.

Der Startvektor werde zufällig gleichverteilt gewählt (also eine zufällige Rundtour, x_1 mit WS'keit $\frac{1}{n}$, dann unabh. x_2 mit WS'keit $\frac{1}{n-1}, \dots$) $\Rightarrow v^{(1)} = \frac{1}{m}(1, \dots, 1)^T$.

Ferner: $v_j^{(k+1)} = \sum_{i=1}^m P_{ij} v_i^{(k)}$, die Summe der WS. dass im k-ten Schritt $x^{(i)}$ gegeben ist multipliziert mit der WS, dass $x^{(j)}$ Nachfolger ist.