Großer Beleg

Effiziente und stabile Berechnung von Varianzen in Markov Ketten.

Maximilian Starke
Student der TU Dresden
Fakultät Informatik

18. Januar 2020

Inhaltsverzeichnis

1 Motivation

hier nur die Aufgabenstellung:

Markov-Ketten mit Kantengewichten dienen der Modellierung und Analyse einer wichtigen Zielgröße in probabilistischen Systemen, z.B. der Dauer eines Verbindungsaufbaus in Computernetzwerken. Im Gegensatz zum Erwartungswert der akkumulierten Gewichte bis zum Erreichen eines Zielzustandes wurden die Schwankungen der Zielgröße in der wissenschaftlichen Literatur kaum untersucht. Dabei sind stark volatile Zielgrößen besonders in sicherheitskritischen Systemen zu vermeiden, was wiederum einen stabilen Algorithmus für ihre Berechnung erfordert.

Nachdem die grundlegende Terminologie ausgearbeitet wurde, ist die Kernaufgabe der Arbeit die formale Erarbeitung und Analyse eines Algorithmus zur Berechnung der Varianz von akkumulierten Gewichten auf Basis von [1]. Der Algorithmus soll implementiert und zur experimentellen Analyse von anwendungsorientierten Beispielen herangezogen werden. Als weitergehende Forschungsfrage soll außerdem untersucht werden, ob das Verfahren auch zur Berechnung von Kovarianzen geeignet ist oder inwieweit der Algorithmus auf Markov-Entscheidungsprozesse mit Kantengewichten verallgemeinert werden kann.

[1] T. Verhoeff, Reward Variance in Markov Chains: A Calculational Approach

META: to do...
* was gibt es an bisheriger forschung
* praktischer Nutzen
* Thema und Ziel dieser arbeit
* Vorabsummary, was den Leser erwartet

2 Grundlagen der Wahrscheinlichkeitstheorie

META: /theoretische/ Grundlagen -¿ Wahrscheinlichkeitstheorie -¿ Markovketten -¿ MDPs bzw. unendliche summen für Beispiel

Definition 2.1 (Wahrscheinlichkeitsraum) Ein Paar (Ω, P) heißt genau dann Wahrscheinlichkeitsraum, wenn Ω eine Menge ist und $P : \Omega \to [0, 1]$ eine Funktion mit

$$\sum_{\omega \in \Omega} P(\omega) = 1. \tag{1}$$

Wir nennen Ω in diesem Zusammenhang auch Ergebnismenge und P eine Wahrscheinlichkeitsverteilung auf Ω

Definition 2.2 (Zufallsvariable) Sei (Ω, P) ein Wahrscheinlichkeitsraum. Eine Funktion $X : \Omega \to \mathbb{R}$ heißt Zufallsvariable auf (Ω, P). Wir nennen X auch kurz Zufallsvariable, falls der Kontext (Ω, P) klar ist.

Definition 2.3 (Erwartungswert) Sei (Ω, P) ein Wahrscheinlichkeitsraum und X eine Zufallsvariable auf (Ω, P).

$$\mathcal{E}_{(\Omega,P)}(X) := \sum_{\omega\in\Omega} P(\omega) \cdot X(\omega) \tag{2}$$

ist der Erwartungswert von X auf (Ω, P). Sollte der Kontext (Ω, P) klar sein, schreiben wir auch kurz $\mathcal{E}(X)$.

Es lässt sich leicht die folgende Linearität des Erwartungswertes beobachten. Seien X, Y Zufallsvariablen, $c, d \in \mathbb{R}$. Dann gilt:

$$\begin{aligned}
\mathcal{E}(X + cY + d) &= \sum_{\omega\in\Omega} P(\omega) \cdot (X + cY + d)(\omega) \\
&= \sum_{\omega\in\Omega} P(\omega) \cdot X(\omega) + c \cdot P(\omega) \cdot Y(\omega) + d \cdot P(\omega) \\
&= \mathcal{E}(X) + c\mathcal{E}(Y) + d \qquad (3)
\end{aligned}$$

Dabei ist $X + cY + d$ die üblich Notation für die Zufallsvariable gegeben durch $Z : \Omega \to \mathbb{R} : \omega \mapsto X(\omega) + c \cdot Y(\omega) + d$.

Definition 2.4 (Varianz) Sei (Ω, P) ein Wahrscheinlichkeitsraum und X eine Zufallsvariable auf (Ω, P).

$$\mathcal{V}_{(\Omega,P)}(X) := \mathcal{E}_{(\Omega,P)}\Big(\big(X - \mathcal{E}_{(\Omega,P)}(X)\big)^2\Big) \tag{4}$$

ist die Varianz von X auf (Ω, P). Sollte der Kontext (Ω, P) klar sein, schreiben wir auch kurz $\mathcal{V}(X)$.

Definition 2.5 (Kovarianz) Seien (Ω, P) ein Wahrscheinlichkeitsraum und X, Y Zufallsvariablen auf (Ω, P). Die Kovarianz $\mathcal{C}_{(\Omega,P)}(X, Y)$ ist definiert durch:

$$\mathcal{C}_{(\Omega,P)}(X, Y) := \mathcal{E}\Big(\big(X - \mathcal{E}(X)\big)\big(Y - \mathcal{E}(Y)\big)\Big) \tag{5}$$

Offenbar ist die Varianz ein Spezialfall der Kovarianz, denn aus den Definitionen folgt unmittelbar $\mathcal{V}_{(\Omega,P)}(X) = \mathcal{C}_{(\Omega,P)}(X, X)$. Wir betrachten nun, wie sich Kovarianzen als Erwartungswerte beschreiben lassen.

Lemma 2.6 *Seien X und Y Zufallsvariablen. Dann gilt:*

$$\mathcal{C}(X, Y) = \mathcal{E}(XY) - \mathcal{E}(X)\mathcal{E}(Y) \tag{6}$$

Beweis.

$$\begin{aligned}
\mathcal{C}(X, Y) &= \mathcal{E}\big(\,(X - \mathcal{E}(X))\,(Y - \mathcal{E}(Y))\,\big) && \text{(Definition 2.5)} \\
&= \mathcal{E}\big(XY - X\mathcal{E}(Y) - Y\mathcal{E}(E) + \mathcal{E}(X)\mathcal{E}(Y)\big) \\
&= \mathcal{E}(XY) - 2\mathcal{E}(X)\mathcal{E}(Y) + \mathcal{E}(X)\mathcal{E}(Y) && \text{(Linearität (3))} \\
&= \mathcal{E}(XY) - \mathcal{E}(X)\mathcal{E}(Y)
\end{aligned}$$

□

Der Spezialfall von Lemma 2.6 für Varianzen lautet dann wiefolgt.

Korollar 2.7 *Sei X eine Zufallsvariable. Dann gilt:*

$$\mathcal{V}(X) = \mathcal{E}(X^2) - \mathcal{E}(X)^2 \tag{7}$$

Analog zu (3) lässt sich für Varianzen folgende Beziehung feststellen:

Lemma 2.8 *Sei X eine Zufallsvariable, $c, d \in \mathbb{R}$. Dann gilt:*

$$\mathcal{V}(cX + d) = c^2 \mathcal{V}(X) \tag{8}$$

Beweis. tbd (steht im Paper) □

META: das korollar wird nicht benutzt....

Korollar 2.9 *Sei X eine Zufallsvariable und $c \in \mathbb{R}$. Dann gilt:*

$$\mathcal{E}\left((c + X)^2\right) = (c + \mathcal{E}(X))^2 + \mathcal{V}(X) \tag{9}$$

Beweis. Aus Korollar 2.7 folgt unmittelbar $\mathcal{V}(X+c) = \mathcal{E}\left((X + c)^2\right) - (\mathcal{E}(X + c))^2$. Dabei kann die linke Seite nach Lemma 2.8 vereifacht werden zu $\mathcal{V}(X + c) = \mathcal{V}(X)$. Der Subtrahend $(\mathcal{E}(X + c))^2$ lässt sich aufgrund der Linearität (Gleichung 3) auch als $(\mathcal{E}(X) + c)^2$ schreiben. □

3 Markovketten

Definition 3.1 (Markovkette) Eine Markovkette ist ein Tupel $M = (Q, P, I)$ mit den Eigenschaften

(a) Q ist eine Menge.

(b) $P : Q \times Q \to [0, 1]$ mit der Eigenschaft $\forall q \in Q : \sum_{q' \in Q} P(q, q') = 1$

(c) $I : Q \to [0, 1]$ ist eine Wahrscheinlichkeitsverteilung auf Q.

Die Bilder von P nennen wir auch Transitionswahrscheinlichkeiten.

Definition 3.2 (Pfad) Sei $M = (Q, P, I)$ eine Markovkette. Die Menge aller (endlichen) Pfade in M ist definiert durch

$$\mathrm{Paths}(M) := \{p \in Q^k \mid k \in \mathbb{N}_+\} \tag{10}$$

Wir nennen einen Pfad $p \in \mathrm{Paths}(M)$ echt, wenn zusätzlich $\mathrm{real}_M(p)$ gilt mit:

$$\mathrm{real}_M(p) := \forall 0 \le i < |p| - 1 : P(p_i, p_{i+1}) > 0 \tag{11}$$

Die Wahrscheinlichkeit $\mathrm{P}(p)$ eines Pfades ist gegeben durch

$$\mathrm{P}(p) := \prod_{i=0}^{|p|-2} P(p_i, p_{i+1}) \tag{12}$$

Insbesondere ist $\mathrm{P}(p) = 1$ für alle Pfade p mit $|p| = 1$ und ein Pfad p ist genau dann echt, wenn $\mathrm{P}(p) > 0$, d.h. wenn p tatsächlich ein Streckenzug im zugrundeliegenden gerichteten Graph von M ist.

Für $|p| = 2$ entspricht die Wahrscheinlichkeit $\mathrm{P}(p)$ genau der Wahrscheinlichkeit gegeben durch die Funktion P der Transitionswahrscheinlichkeiten aus der Markovkette. Daher gilt $P \subseteq \mathrm{P}$, P ergibt sich als eindeutige Fortsetzung von P und wir schreiben im Folgenden einfach P.

Definition 3.3 (Gewichtsfunktion) Sei $M = (Q, P, I)$ eine Markovkette. Eine Gewichtsfunktion auf M ist eine Abbildung

$$R : Q \times Q \to \mathbb{R}. \tag{13}$$

Offenbar gibt es analog zur Wahrscheinlichkeit P eine eindeutige Fortsetzung für eine Gewichtsfunktion R, gegeben durch Aufsummierung aller Kantengewichte entlang von Pfaden:

$$\mathrm{R} : \bigcup_{k \in \mathbb{N}_+} Q^k \to \mathbb{R} : p \mapsto \sum_{i=0}^{|p|-2} R(p_i, p_{i+1}) \tag{14}$$

Wir schreiben im Folgenden einfach R.

Definition 3.4 Sei $M = (Q, P, I)$ eine Markovkette, $s \in Q$ ein Startzustand und $A \subseteq Q$ eine Menge von Zielzuständen. Wir bezeichnen die Menge der Pfade, welche in s starten und in A enden, jedoch A nicht zwischenzeitlich schon erreichen mit:

$$\mathrm{Paths}_{s \to A}(M) = \{p \in \mathrm{Paths}(M) \mid p_0 = s \wedge p_{|p|-1} \in A \wedge \forall i < |p| - 1 : p_i \notin A\} \tag{15}$$

META: [Literaturverweis einfügen:] Wenn A von allen $q \in Q$ erreichbar, dann ...

Wir beobachten, dass unter dieser Voraussetzung $(\mathrm{Paths}_{s \to A}(M), P)$ ein Wahrscheinlichkeitsraum ist und R eine Zufallsvariable auf M.

Definition 3.5 Sei $M = (Q, P, I)$ eine endliche Markovkette, $A \subseteq Q$ eine Zielmenge. Dann bezeichnen wir mit $P_{\to A}$ die folgende Matrix:

$$P_{\to A} : Q^2 \to [0, 1] : \begin{cases} P(s, t) & \text{falls } s \notin A \\ 0 & \text{falls } s \in A \end{cases} \tag{16}$$

Diese Modifikation entspricht dem Entfernen von allen ausgehenden Kanten von Zuständen $a \in A$.

Satz 3.6 *Sei $M = (Q, P, I)$ eine endliche Markovkette, $A \subseteq Q$ eine Zielmenge, die von jedem Knoten aus erreichbar ist ($\forall s \in Q : \exists p \in \mathrm{Paths}_{s \to A}(M) : \mathrm{real}_M(p)$). Dann ist die Matrix $M := P_{\to A} - \mathbb{1}$ invertierbar.*

Beweis. Aus der linearen Algebra ist bekannt, dass M genau dann invertierbar ist, wenn $Mx = 0$ nur die Lösung $x = 0$ hat. Nehmen wir also an, $x \in R^Q$ mit $x \neq 0$ erfüllt $Mx = 0$. Dann gilt auch $\forall k \in \mathbb{R} : M(kx) = 0$. Wählen wir ein geeignetes k, so erhalten wir eine Lösung z mit einem positiven Eintrag, d.h. $Mz = 0$ und $\exists q \in Q : z_q > 0$. Sei $E \subseteq Q$ die Indexmenge aller maximalen Einträge von z, also $e \in E :\Leftrightarrow \forall q \in Q z_q \leq z_e$. Sei $s \in A$. Dann folgt aus $Mx_s = 0_s$, dass $\sum_{q \in Q} (\widehat{P} - \mathbb{1})_{(s,q)} \cdot z_q = \sum_{q \in Q} -\mathbb{1}_{(s,q)} \cdot z_q = 0$ und damit $z_s = 0$.

Sei $s \in E$. Dann ist $z_s > 0$ und daher $s \notin A$. Aus $Mx_s = 0_s$ folgt dann $\sum_{q \in Q} (\widehat{P} - \mathbb{1})_{(s,q)} \cdot z_q = 0$. Damit bekommen wir $\sum_{q \in Q} P_{(s,q)} \cdot z_q = z_s$. Da nach Definition 3.1 $q \mapsto P(s, q)$ eine Wahrscheinlichkeitsverteilung auf Q ist, folgt schließlich $\forall q \in Q : P(s, q) > 0 \Rightarrow z_q = z_s$. D.h. für jeder Nachfolgerknoten $t \in Q$ von $s \in E$ in der Markovkette M liegt selbst wieder in E. Da ein Knoten $a \in A$ von s aus erreichbar ist, gilt auch $a \in E$. Wir erhalten einen Widerspruch. $\square$

4 Formale Herleitung eines Algorithmus

Ziel ist es nun, Algorithmen zu beschreiben und zu analysieren, welche bei gegebener Markovkette M, gegebenem Startzustand s und gegebener Zielzustandsmenge A die Varianz bzw. Kovarianz berechnen. Wir beschränken uns dabei auf den Fall, dass A von jedem Zustand in M aus erreichbar ist.

4.1 Berechnung von Erwartungswerten

Seien $M = (Q, P, I)$ eine Markovkette, $s \in Q$, $\emptyset \neq A \subseteq Q$ und $\forall s \in Q : \exists p \in \mathrm{Paths}_{s \to A}(M) : \mathrm{real}_M(p)$. Sei R eine Gewichtsfunktion auf M. Wir betrachten im Wahrscheinlichkeitsraum $(\mathrm{Paths}_{s \to A}(M), P)$ den Erwartungswert $\mathcal{E}_{(\mathrm{Paths}_{s \to A}(M), P)}(R)$, im Folgenden kurz $\mathcal{E}_s(R)$:

$$\mathcal{E}_s(R) = \sum_{p \in \mathrm{Paths}_{s \to A}(M)} P(p) \cdot R(p) \tag{17}$$

Falls $s \in A$, gilt $|\mathrm{Paths}_{s \to A}(M)| = 1$ mit dem einzigen enthaltenen Pfad $p = (s)$. Dann sind nach Definition $P(p) = 1$ und $R(p) = 0$. Damit erhalten wir:

$$\mathcal{E}_s(R) = 0 \qquad\qquad (\text{falls } s \in A) \tag{18}$$

Falls $s \notin A$, besteht jeder Pfad $p \in \mathrm{Paths}_{s \to A}(M)$ aus mehr als einem Knoten, und wir erhalten:

$$\begin{aligned}
\mathcal{E}_s(R) &= \sum_{p \in \mathrm{Paths}_{s \to A}(M)} P(s, p_1) \cdot P(p_{\leftarrow 1}) \cdot (R(s, p_1) + R(p_{\leftarrow 1})) && (19)\\
&= \sum_{t \in Q} P(s,t) \cdot \sum_{p' \in \mathrm{Paths}_{t \to A}(M)} P(p') \cdot (R(s,t) + R(p')) && (20)\\
&= \sum_{t \in Q} P(s,t) \cdot \left((R(s,t) + \sum_{p' \in \mathrm{Paths}_{t \to A}(M)} P(p') \cdot R(p') \right) && (21)\\
&= \sum_{t \in Q} P(s,t) \cdot (R(s,t) + \mathcal{E}_t(R)) && (22)
\end{aligned}$$

META: TODO: notation erklären $p_{\leftarrow 1}$

Die Gleichungen 18 und 22 geben uns ein System von $|Q|$ linearen Gleichungen in $|Q|$ Variablen:

$$\begin{aligned}
\mu_s &= 0 && (\text{falls } s \in A)\\
\mu_s &= \sum_{t \in Q} P(s,t) \cdot (R(s,t) + \mu_t) && (\text{falls } s \notin A)
\end{aligned} \tag{23}$$

Mit Definition 3.5 ist dieses Gleichungssystem äquivalent zu

$$\mu_s = \sum_{t \in Q} P_{\to A}(s,t) \cdot (R(s,t) + \mu_t), \tag{24}$$

was sich in Matrixschreibweise mit $\mu = (\mu_s)_{s \in Q}$ ausdrücken lässt als:

$$\mu = \left(\sum_{t \in Q} P_{\to A}(s,t) \cdot R(s,t) \right)_{s \in Q} + P_{\to A} \cdot \mu \tag{25}$$

$$\iff \qquad (P_{\to A} - \mathbb{1})\mu = - \left(\sum_{t \in Q} P_{\to A}(s,t) \cdot R(s,t) \right)_{s \in Q} \tag{26}$$

Wir stellen fest, der Vektor $(\mathcal{E}_s(R))_{s \in Q}$ ist die einzige Lösung des Gleichungssystems (26): Zum einen ist gemäß unserer Herleitung ist der Vektor $(\mathcal{E}_s(R))_{s \in Q}$ eine Lösung dieses Gleichungssystems. Zum anderen ist $P_{\to A} - \mathbb{1}$ nach Satz 3.6 invertierbar. Somit ist die Lösung eindeutig.

Mit dem Gleichungssystem (23) und Standardalgorithmen zum Lösen linearer Gleichungssysteme erhalten wir unmittelbar einen Algorithmus zur Berechnung der Erwartungswerte $E_q(R)$ für $q \in Q$.

4.2 Berechnung von Varianzen

Seien $M = (Q, P, I)$ eine Markovkette, $s \in Q$, $\emptyset \neq A \subseteq Q$ und $\forall s \in Q : \exists p \in \mathrm{Paths}_{s \to A}(M) : \mathrm{real}_M(p)$. Sei R eine Gewichtsfunktion auf M. Wir betrachten im Wahrscheinlichkeitsraum $(\mathrm{Paths}_{s \to A}(M), P)$ die Varianz $\mathcal{V}_{(\mathrm{Paths}_{s \to A}(M), P)}(R)$:

$$\mathcal{V}_{(\mathrm{Paths}_{s \to A}(M), P)}(R) = \mathcal{E}_{(\mathrm{Paths}_{s \to A}(M), P)} \left(\left(R - \mathcal{E}_{(\mathrm{Paths}_{s \to A}(M), P)}(R) \right)^2 \right) \tag{27}$$

Wir nutzen auch hier wieder die Kurzschreibweise $\mathcal{V}_s(R) := \mathcal{V}_{(\mathrm{Paths}_{s \to A}(M), P)}(R)$ und erhalten:

$$\mathcal{V}_s(R) = \mathcal{E}_s \left((R - \mathcal{E}_s(R))^2 \right) \tag{28}$$

Falls $s \in A$, gilt $|\mathrm{Paths}_{s \to A}(M)| = 1$ mit dem einzigen enthaltenen Pfad $p = (s)$. Dann sind nach Definition $P(p) = 1$ und $R(p) = 0$ und der Erwartungswert beträgt $\mathcal{E}_s(R) = 0$, wie wir bereits im vorangegangenen Abschnitt gesehen haben. Damit erhalten wir:

$$\mathcal{V}_s(R) = 0 \qquad \text{(falls } s \in A) \tag{29}$$

Falls $s \notin A$, besteht jeder Pfad $p \in \mathrm{Paths}_{s \to A}(M)$ aus mehr als einem

Knoten, und wir erhalten:

$$\mathcal{V}_s(R) = \sum_{p \in \mathrm{Paths}_{s\to A}(M)} P(p) \cdot (R(p) - \mathcal{E}_s(R))^2 \tag{30}$$

$$= \sum_{t\in Q} P(s,t) \cdot \sum_{p' \in \mathrm{Paths}_{t\to A}(M)} P(p') \cdot (R(s,t) + R(p') - \mathcal{E}_s(R))^2 \tag{31}$$

$$= \sum_{t\in Q} P(s,t) \cdot \mathcal{E}_t \left((R + R(s,t) - \mathcal{E}_s(R))^2 \right) \tag{32}$$

$$= \sum_{t\in Q} P(s,t) \cdot \left(\mathcal{V}_t(R) + \left(\mathcal{E}_t\big(R + R(s,t) - \mathcal{E}_s(R)\big) \right)^2 \right) \quad \text{(Korollar (2.7))} \tag{33}$$

$$= \sum_{t\in Q} P(s,t) \cdot \left(\mathcal{V}_t(R) + \big(\mathcal{E}_t(R) + R(s,t) - \mathcal{E}_s(R)\big)^2 \right) \quad \text{(Linearität (3))} \tag{34}$$

Die Gleichungen 29 und 34 geben uns ähnlich wie im Abschnitt über die Erwartungswerte ein System von $|Q|$ linearen Gleichungen in $|Q|$ Variablen:

$$\begin{aligned} \nu_s &= 0 \quad (\text{falls } s \in A) \\ \nu_s &= \sum_{t\in Q} P(s,t) \cdot \left(\nu_t + \big(\mathcal{E}_t(R) + R(s,t) - \mathcal{E}_s(R)\big)^2 \right) \quad (\text{falls } s \notin A) \end{aligned} \tag{35}$$

Sei die Gewichtsfunktion S auf definiert durch $S : Q^2 \to \mathbb{R}_+ : (s,t) \mapsto \big(\mathcal{E}_t(R) + R(s,t) - \mathcal{E}_s(R)\big)^2$ und sei $\nu = (\nu_s)_{s\in Q}$. Analog zu Gleichung 26 bekommen wir durch Anwendung von Definition 3.5 das äquivalente Gleichungsystem

$$(P_{\to A} - \mathbb{1})\nu = -\left(\sum_{t\in Q} P_{\to A}(s,t) \cdot S(s,t) \right)_{s\in Q} \tag{36}$$

Auch hier besitzt das Gleichungsystem eine eindeutige Lösung nach Satz 3.6. Vergleicht man die Gleichungssysteme (26) und (36), dann bekommen wir:

$$\forall q \in Q : \mathcal{V}_q(R) = \mathcal{E}_q(S) \tag{37}$$

Damit haben wir die Berechnung der Varianzen gleichzeitig auf die Berechnung von Erwartungswerten zurückgeführt. Um die Varianzen $(\mathcal{V}_q(R))_{q\in Q}$ in einer Markovkette zu berechnen, genügt es die Erwartungswerte $(\mathcal{E}_q(R))_{q\in Q}$ im ersten Schritt und $(\mathcal{E}_q(S))_{q\in Q}$ im zweiten Schritt zu berechnen. Es ist sogar noch eine Optimierung möglich, indem zuerst die inverse Matrix $(P_{\to A} - \mathbb{1})^{-1}$ für sich berechnet wird. Dann müssen für das Lösen beider Gleichungssysteme jeweils nur eine Matrixmultiplikation ausgeführt werden.

4.3 Berechnung von Kovarianzen

Seien wieder $M = (Q, P, I)$ eine Markovkette, $s \in Q$, $\emptyset \neq A \subseteq Q$ und $\forall s \in Q : \exists p \in \mathrm{Paths}_{s\to A}(M) : \mathrm{real}_M(p)$. Seien nun X, Y Gewichtsfunktionen auf M. Wir betrachten im Wahrscheinlichkeitsraum $(\mathrm{Paths}_{s\to A}(M), P)$ die Kovarianz $\mathcal{C}_s(X,Y) := \mathcal{C}_{(\mathrm{Paths}_{s\to A}(M),P)}(X,Y)$:

$$\mathcal{C}_s(X,Y) = \mathcal{E}_s\big(\,(X - \mathcal{E}_s(X))\,(Y - \mathcal{E}_s(Y))\,\big) \tag{38}$$

Falls $s \in A$, gilt $|\text{Paths}_{s\to A}(M)| = 1$ mit dem einzigen enthaltenen Pfad $p = (s)$. Dann sind nach Definition $P(p) = 1$ und $R(p) = 0$ und die Erwartungswerte betragen $\mathcal{E}_s(X) = \mathcal{E}_s(Y) = 0$. Damit erhalten wir:

$$\mathcal{C}_s(X,Y) = 0 \qquad \text{(falls } s \in A\text{)} \tag{39}$$

Falls $s \notin A$, besteht jeder Pfad $p \in \text{Paths}_{s\to A}(M)$ aus mehr als einem Knoten, und wir erhalten:

$$\mathcal{V}_s(R) = \sum_{p\in\text{Paths}_{s\to A}(M)} P(p)\cdot\big(\,(X(p) - \mathcal{E}_s(X))\,(Y(p) - \mathcal{E}_s(Y))\,\big) \tag{40}$$

$$= \sum_{t\in Q} P(s,t)\cdot \sum_{p'\in\text{Paths}_{t\to A}(M)} P(p')\cdot\big((X(s,t) + X(p') - \mathcal{E}_s(X))(Y(s,t) + Y(p') - \mathcal{E}_s(Y))\big) \tag{41}$$

$$= \sum_{t\in Q} P(s,t)\cdot \mathcal{E}_t\Big(\big(X + X(s,t) - \mathcal{E}_s(X)\big)\big(Y + Y(s,t) - \mathcal{E}_s(Y)\big)\Big) \tag{42}$$

$$= \sum_{t\in Q} P(s,t)\cdot \begin{pmatrix} \mathcal{E}_t(XY) & +\mathcal{E}_t(X)(Y(s,t) - \mathcal{E}_s(Y)) \\ +\,\mathcal{E}_t(Y)(X(s,t) - \mathcal{E}_s(X)) & +(X(s,t) - \mathcal{E}_s(X))(Y(s,t) - \mathcal{E}_s(Y)) \end{pmatrix} \tag{43}$$

$$= \sum_{t\in Q} P(s,t)\cdot \begin{pmatrix} \mathcal{E}_t(XY) & -\mathcal{E}_t(X)\mathcal{E}_t(Y) \\ +\,\mathcal{E}_t(X)\mathcal{E}_t(Y) & +\mathcal{E}_t(X)(Y(s,t) - \mathcal{E}_s(Y)) \\ +\,\mathcal{E}_t(Y)(X(s,t) - \mathcal{E}_s(X)) & +(X(s,t) - \mathcal{E}_s(X))(Y(s,t) - \mathcal{E}_s(Y)) \end{pmatrix} \tag{44}$$

$$= \sum_{t\in Q} P(s,t)\cdot \Big(\mathcal{C}_t(X,Y) + \big(\mathcal{E}_t(X) + X(s,t) - \mathcal{E}_s(X)\big)\big(\mathcal{E}_t(Y) + Y(s,t) - \mathcal{E}_s(Y)\big)\Big) \tag{45}$$

Die Gleichungen 39 und 45 geben uns ähnlich wie im Abschnitt über die Erwartungswerte ein System von $|Q|$ linearen Gleichungen in $|Q|$ Variablen:

$$\begin{aligned} c_s &= 0 && \text{(falls } s \in A\text{)} \\ c_s &= \sum_{t\in Q} P(s,t)\cdot\Big(\big(\mathcal{E}_t(X) + X(s,t) - \mathcal{E}_s(X)\big)\big(\mathcal{E}_t(Y) + Y(s,t) - \mathcal{E}_s(Y)\big) + c_t\Big) && \text{(falls } s \notin A\text{)} \end{aligned} \tag{46}$$

Sei die Gewichtsfunktion S auf M definiert durch $S : Q^2 \to \mathbb{R} : (s,t) \mapsto \big(\mathcal{E}_t(X) + X(s,t) - \mathcal{E}_s(X)\big)\big(\mathcal{E}_t(Y) + Y(s,t) - \mathcal{E}_s(Y)\big)$ und sei $c = (c_s)_{s\in Q}$. Analog zu Gleichung 26 bekommen wir durch Anwendung von Definition 3.5 das äquivalente Gleichungsystem

$$(P_{\to A} - \mathbb{1})c = -\left(\sum_{t\in Q} P_{\to A}(s,t)\cdot S(s,t)\right)_{s\in Q} \tag{47}$$

Auch hier besitzt das Gleichungsystem eine eindeutige Lösung nach Satz 3.6. Vergleicht man die Gleichungssysteme (26) und (47), dann bekommen wir:

$$\forall q \in Q : \mathcal{C}_q(X,Y) = \mathcal{E}_q(S) \tag{48}$$

Damit haben wir die Berechnung der Kovarianzen gleichzeitig auf die Berechnung von Erwartungswerten zurückgeführt. Um die Kovarianzen $(\mathcal{C}_q(X,Y))_{q\in Q}$ in einer Markovkette zu berechnen, genügt es die Erwartungswerte $(\mathcal{E}_q(X))_{q\in Q}$ sowie $(\mathcal{E}_q(Y))_{q\in Q}$ im ersten Schritt und $(\mathcal{E}_q(S))_{q\in Q}$ im zweiten Schritt zu berechnen. Es ist wie bei den Varianzen die Optimierung möglich, zuerst die inverse Matrix $(P_{\to A} - \mathbb{1})^{-1}$ zu berechnen, um danach nur noch Matrixmultiplikationen ausführen zu müssen.

5 Performancemessung am praktischen Beispiel

5.1 Herman's Selbststabilisierungsalgorithmus

Hier verweis einfügen

Um die beschriebene Berechnung von Erwartungswerten, Varianzen und Kovarianzen praktisch zu demonstrieren soll uns hier ein Beispiel dienen, welches auch in den *PRISM Case Studies* zu finden ist.

META: Anmerkung: Die Berechnung der Varianz ist mittels 2.7 kann zu unschönen Fehlern mit floatingpoint Standard führen...

use https://github.com/ddemidov/amgcl
Eigen library?

Literatur

[1] *bla bla* https://de.wikipedia.org/wiki/BibTeX