

Relationen bedingter Wahrscheinlichkeiten

1 Projektive Räume und homogene Koordinaten

Sei V ein \mathbb{K} -Vektorraum der Dimension $n + 1$. Identifiziert man in diesem Vektorraum jeweils die Elemente jedes 1-dimensionalen Unterraums außer der 0, so erhält man einen projektiven Raum $\mathbb{P}(V)$ der Dimension n . Es gilt also

$$\mathbb{P}(V) := (V \setminus \{0\}) / \sim,$$

wobei für a, b aus V gilt:

$$a \sim b \Leftrightarrow \exists \lambda \in \mathbb{K} \setminus \{0\} : a = \lambda \cdot b.$$

Im Fall $V = \mathbb{K}^{n+1}$ schreibt man $\mathbb{P}(V)$ als \mathbb{P}^n und die Elemente aus \mathbb{P}^n in homogenen Koordinaten $(v_0 : \dots : v_n)$. Offenbar ist die Darstellung eines solchen Punktes in homogenen Koordinaten nicht eindeutig, da

$$\lambda \cdot (v_0 : \dots : v_n) = (\lambda v_0 : \dots : \lambda v_n) = (v_0 : \dots : v_n), \forall \lambda \in \mathbb{K} \setminus \{0\}$$

Mit der injektiven Abbildung

$$\begin{aligned} \pi : \quad \mathbb{K}^n &\longrightarrow \mathbb{P}^n \\ (a_1, \dots, a_n) &\longmapsto (1 : a_1 : \dots : a_n) \end{aligned}$$

kann man \mathbb{K}^n als Teilmenge des \mathbb{P}^n verstehen. Es gilt $\text{Bild}(\pi) = \{a \in \mathbb{P}^n : a_0 \neq 0\}$. Anschaulich erhält man den \mathbb{P}^n indem man den \mathbb{K}^n um eine unendlich ferne Hyperebene erweitert; die Punkte $(1 : \dots)$ repräsentieren dabei die affinen Punkte, die Punkte $(0 : \dots)$ repräsentieren die unendlich fernen Punkte, wobei der hintere Teil der Koordinaten die Richtung des unendlich fernen Punktes angibt.

In der Abbildung 1 sind die Konstruktionen des \mathbb{P}^1 und des \mathbb{P}^2 veranschaulicht. Der Halbkreis im linken Bild ist isomorph zu $\mathbb{R}^1 \cup \{\infty\} \cong \mathbb{P}^1$. Im rechten Bild sieht man die Einbettung von \mathbb{R}^2 in $\mathbb{P}(\mathbb{R}^3) = \mathbb{P}^2$. Die Geraden, die die eingezeichnete Ebene schneiden, entsprechen im \mathbb{R}^2 ihren Schnittpunkten mit dieser Ebene. Die Geraden, die parallel zur Ebene verlaufen, entsprechen den unendlichen fernen Punkten, genauer sogar dem \mathbb{P}^1 , denn es gilt $\mathbb{P}^2 \cong \mathbb{R}^2 \cup \mathbb{P}^1$. Interpretiert man die im Bild vertikale Koordinatenachse als die x_1 -Achse, so sieht man anschaulich, warum die Punkte mit $x_1 = 0$ den unendlich fernen

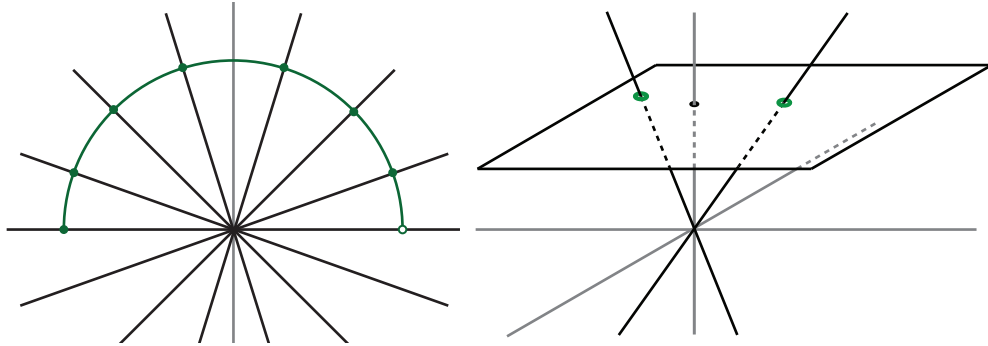


Abbildung 1: Konstruktion des \mathbb{P}^1 und des \mathbb{P}^2 über \mathbb{R}

Punkten und die Punkte mit $x_1 = 1$ den affinen Punkten entsprechen: Die Geraden, die parallel zur Ebene verlaufen, werden genau von solchen Vektoren repräsentiert, für die $x_1 = 0$ gilt.

Die Wahl von $i = 1$ ist hier willkürlich, ebenso kann jede beliebige Hyperebene des \mathbb{P}^n als unendlich ferne Hyperebene und das Komplement als \mathbb{K}^n aufgefasst werden. In der Konstruktion des \mathbb{P}^2 in Abbildung 1 entspricht dies einer veränderten Wahl der Hyperebene, dessen Schnittpunkte mit den Geraden als affine Punkte aufgefasst werden. Im Fall $n \leq 3$ kann man so verschiedene Karten aufzeichnen und auch Schnitte im unendlich Fernen gut visualisieren.

Betrachtet man Nullstellenmengen von Polynomen im projektiven Raum, so stößt man zunächst auf ein Problem der Wohldefiniertheit: Für Polynome $f \in \mathbb{K}[x_0, \dots, x_n]$ mit einer Nullstelle $a = (a_0, \dots, a_n)$ gilt in der Regel nicht $f(\lambda \cdot a) = 0$; aber es ist $a = \lambda \cdot a$, wenn a ein Punkt im projektiven Raum ist. Dieses Problem kann man durch den Übergang auf homogene Polynome umgehen, d.h. Polynome, die in jedem Monom denselben Grad haben. Davon erzeugte Ideale nennt man homogene Ideale und auf diese Weise kann man analog zum Affinen eine Theorie der projektiven Varietäten aufbauen.

Mehr über projektive Geometrie und der Theorie der affinen und projektiven Varietäten ist in [1] zu finden. Im Folgenden soll, falls nicht anders erwähnt, \mathbb{C} als Grundkörper vorausgesetzt werden.

2 Bedingte Wahrscheinlichkeitsverteilungen

Sei $\Omega = [m]$ ein endlicher Ereignisraum, $\mathcal{E} \subset 2^\Omega$ wobei für alle $I \in \mathcal{E}$ gilt $|I| \geq 2$. $\mathbb{C}[\mathcal{E}]$ bezeichne die Ereignisalgebra, der Polynomring mit Unbestimmten $p_{i|I}$ für alle $I \in \mathcal{E}$ und $i \in I$, d.h. eine Unbekannte für jede elementare bedingte Wahrscheinlichkeit. Es bezeichne weiter $|\mathcal{E}| = \sum_{I \in \mathcal{E}} |I|$ die Anzahl der Variablen in $\mathbb{C}[\mathcal{E}]$; p_i für $p_{i|[m]}$, falls $[m] \in \mathcal{E}$.

Ein Punkt $(p_1, \dots, p_m) \in \mathbb{R}_{\geq}^m$ mit $\sum_{i=1}^m p_i = 1$ repräsentiert eine Wahrscheinlichkeitsverteilung auf Ω . Gilt $p_j > 0$ für alle j , so ist die bedingte Wahrschein-

lichkeit von i gegeben $I \ni i$:

$$p_{i|I} = \frac{p_i}{\sum_{j \in I} p_j}.$$

Allgemeiner folgende

Definition 2.1 Eine bedingte Wahrscheinlichkeitsverteilung für \mathcal{E} ist ein Punkt $(p_{i|I} : i \in I \in \mathcal{E}) \in \mathbb{R}_{\geq 0}^{|\mathcal{E}|}$, so dass für alle $J, K \in \mathcal{E}$ mit $J \subset K$ gilt

(i) $\sum_{i \in J} p_{i|J} = 1$,

(ii) für alle $i \in J$ gilt $p_{i|K} = p_{i|J} \sum_{j \in J} p_{j|K}$.

Dies ist die verallgemeinerte Version der obigen Formel, welche man aus der Definition 2.1 erhält, wenn man K zu $[m]$ und J zu I spezialisiert und $\sum_{i \in I} p_i \neq 0$ voraussetzt. Wenn allerdings $\sum_{i \in I} p_i = 0$ gilt, so erfüllt der gesamte Wahrscheinlichkeitssimplex $\Delta_J := \{(p_{j|J})_{j \in J} : p_{j|J} \geq 0, \sum_{j \in J} p_{j|J} = 1\}$ diese Definition.

2.1 Eine projektive Sicht von Wahrscheinlichkeit

Ausgehend von Definition 2.1 erscheint es naheliegend, nicht einfache Punkte $(p_1, \dots, p_m) \in \mathbb{R}_{\geq 0}^m$ mit $\sum_{i=1}^m p_i = 1$ zu betrachten sondern diese Punkte stattdessen als Elemente des $\mathbb{P}^{m-1} = \mathbb{P}(\mathbb{C}^m)$ in homogenen Koordinaten zu verstehen. So eliminiert man die Bedingung $\sum_{i=1}^m p_i = 1$. Den Begriff des Wahrscheinlichkeitssimplexes kann man auf zwei Arten mit diesem komplex-projektiven Raum zusammenführen:

Der offensichtlichste Weg ist die Restriktion, indem man den Wahrscheinlichkeitssimplex Δ_{m-1} mit dem reellen positiven Teil von $\{y = (y_1 : \dots : y_m) \in \mathbb{P}^{m-1} : \sum_{i=1}^m y_i \neq 0\}$ identifiziert. In Abbildung 2 ist dies für den Fall $m = 3$ illustriert. Hier ist die Hyperebene $V(y_1 + y_2 + y_3)$ die unendlich ferne.

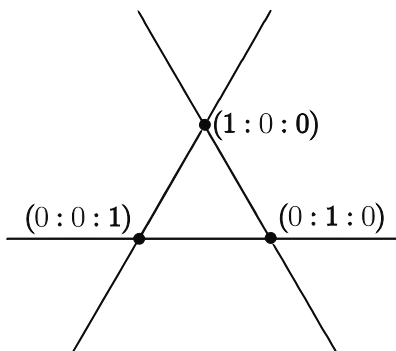


Abbildung 2: projektiver Simplex

Aufgrund der fehlenden algebraischen Struktur erweist sich dies aber als weniger günstig als der alternative Weg, \mathbb{P}^{m-1} mittels der Abbildung

$$\begin{aligned} \mu : \quad \mathbb{P}^{m-1} &\longrightarrow \Delta_{m-1} \\ (y_1 : \dots : y_m) &\longmapsto \frac{1}{\sum_i |y_i|} |y_i| e_i \end{aligned}$$

surjektiv auf den Wahrscheinlichkeitssimplex Δ_{m-1} zu projizieren. Auf seinem Bild ist μ die Identität, somit kann man nun beliebige Punkte des komplex projektiven Raums als Wahrscheinlichkeitsverteilungen interpretieren. So lassen sich nun die Relationen bedingter Wahrscheinlichkeiten mit der Struktur projektiver Varietäten beschreiben.

2.2 Homogene bedingte Wahrscheinlichkeit

Die Definition 2.1 von bedingter Wahrscheinlichkeit soll im Folgenden so verallgemeinert werden, dass die Bedingung, eine bedingte Wahrscheinlichkeitsverteilung zu sein, mit homogenen Polynomen ausgedrückt werden kann. Damit kann die Menge der bedingten Wahrscheinlichkeitsverteilungen als projektive Varietät verstanden werden.

Eine bedingte Wahrscheinlichkeitsverteilung wird nun repräsentiert durch einen Punkt im kartesischen Produkt projektiver Räume. Dieses Produkt besteht aus Faktoren $\mathbb{P}^{|I|-1}$, mit jeweils einem Faktor für jedes $I \in \mathcal{E}$. Jeder dieser Faktoren $\mathbb{P}^{|I|-1}$ enthält wiederum Punkte in homogenen Koordinaten $(p_{i_0|I} : \dots : p_{i_{|I|}|I})$.

Definition 2.2 Eine projektive bedingte Wahrscheinlichkeitsverteilung auf \mathcal{E} ist ein Punkt $p = ((p_{i_0|I} : \dots : p_{i_{|I|}|I}), I \in \mathcal{E}) \in \prod_{I \in \mathcal{E}} \mathbb{P}^{|I|-1}$ so dass für alle $J, K \in \mathcal{E}$ und $i \in J \subset K$ gilt

$$\left(\sum_{j \in J} p_{j|J} \right) p_{i|K} = p_{i|J} \left(\sum_{j \in J} p_{j|K} \right)$$

Für wirkliche bedingte Wahrscheinlichkeitsverteilungen ist $\sum_{j \in J} p_{j|J} = 1$ und dann steht im Wesentlichen die Definition 2.1 da.

Beispiel 2.3 Sei $m = 4$, $\mathcal{E} = \{123, 234, 23\}$. Wähle

$$p = \left(\left(\frac{1}{6} : \frac{2}{6} : \frac{3}{6} \right), \left(\frac{2}{9} : \frac{3}{9} : \frac{4}{9} \right), \left(\frac{2}{5} : \frac{3}{5} \right) \right) \in \mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^1.$$

Nachrechnen zeigt, dass die Gleichungen aus der obigen Definition erfüllt sind, z.B. für $K = 123$, $J = 23$, $i = 2$:

$$\begin{aligned} \left(\sum_{j=2,3} p_{j|23} \right) \cdot p_{2|123} &= p_{2|23} \cdot \left(\sum_{j=2,3} p_{j|123} \right) \\ \Leftrightarrow \quad 1 \cdot \frac{2}{6} &= \frac{2}{5} \cdot \left(\frac{2}{6} + \frac{3}{6} \right) \end{aligned}$$

Bei selbigem \mathcal{E} kann man auch für $q = \left(\left(\frac{1}{6} : \frac{2}{6} : \frac{3}{6} \right), \left(\frac{2}{9}i : \frac{3}{9}i : \frac{4}{9}i \right), \left(-\frac{2}{5} : -\frac{3}{5} \right) \right)$ die Korrektheit der obigen Gleichungen nachrechnen. Es gilt $\mu(q) = p$.

Aus der Definition 2.2 leitet sich das folgende Ideal in der Ereignisalgebra $\mathbb{C}[\mathcal{E}]$ ab:

$$J_{\mathcal{E}} = \langle (\sum_{j \in J} p_{j|J}) p_{i|K} - p_{i|J} (\sum_{j \in J} p_{j|K}) : J, K \in \mathcal{E}, i \in J \subset K \rangle.$$

Dieses Ideal enthält alle Polynombedingungen, die ein Punkt $p \in \prod_{I \in \mathcal{E}} \mathbb{P}^{|I|-1}$ erfüllen muss, damit er eine projektive bedingte Wahrscheinlichkeitsverteilung ist. Insbesondere erfüllt jede wirkliche bedingte Wahrscheinlichkeitsverteilung diese Bedingungen.

Folgende Abkürzungen sollen noch vereinbart werden: es sei $p_{J|J} := \sum_{j \in J} p_{j|J}$. Für wirkliche Wahrscheinlichkeitsverteilungen gilt natürlich $p_{J|J} = 1$, doch $p_{J|J}$ soll hier als Linearform in $\mathbb{C}[\mathcal{E}]$ verstanden werden. Es sei weiter $\alpha_{\mathcal{E}} := \prod_{i \in I \in \mathcal{E}} p_{i|I}$, $\beta_{\mathcal{E}} := \prod_{I \in \mathcal{E}} p_{I|I}$. Die Saturierung eines Ideals I bzgl. eines Polynoms f ist

$$\begin{aligned} (I : f^{\infty}) &:= \langle g \in \mathbb{C}[\mathcal{E}] \mid \exists n \in \mathbb{N} : f^n g \in I \rangle \\ &= \{g \in \mathbb{C}[\mathcal{E}] \mid \exists n \in \mathbb{N} : f^n g \in I\}. \end{aligned}$$

Für den Fall $[m] \in \mathcal{E}$ definiere

$$I_{\mathcal{E}} := (J_{\mathcal{E}} : (\alpha_{\mathcal{E}} \beta_{\mathcal{E}})^{\infty}).$$

Im Falle $[m] \notin \mathcal{E}$, setze $\mathcal{E}' := \mathcal{E} \cup [m]$ und $I_{\mathcal{E}} := I_{\mathcal{E}'} \cap \mathbb{C}[\mathcal{E}]$.

Es gilt $I_{\mathcal{E}} \supseteq J_{\mathcal{E}}$. Die Erweiterung des Ideals $J_{\mathcal{E}}$ auf seine Saturierung verkleinert zunächst ggf. die Nullstellenmenge, d.h. es könnten nun Punkte aus der Nullstellenmenge verschwinden, die aber eigentlich eine projektive bedingte Wahrscheinlichkeitsverteilung im Sinne der Definition 2.2 sind. Doch die Definition von $\alpha_{\mathcal{E}}$ und $\beta_{\mathcal{E}}$ sorgt dafür, dass eben nur solche Wahrscheinlichkeitsverteilungen verschwinden, wo Einzelereignisse i oder ganze Ereignisse I unmöglich sind, und durch Reduzierung von \mathcal{E} und Ω werden diese dann irrelevant.

3 Eine universelle Gröbnerbasis für bedingte Wahrscheinlichkeitsverteilungen

Die universelle Gröbnerbasis von $I_{\mathcal{E}}$ wird aus den Zyklen eines benannten bipartiten Graphen $G(\mathcal{E})$ hervor gehen. Dieser wird folgendermaßen definiert:

Ecken: Eine Ecke u_I für jedes $I \in \mathcal{E}$ sowie eine Ecke v_i für jedes $i \in \cup_{I \in \mathcal{E}} I$.

Kanten: Eine gerichtete Ecke $u_I \rightarrow v_i$ für jedes $I \in \mathcal{E}$ und $i \in I$.

Kantenbezeichnungen: Die Ecke $u_I \rightarrow v_i$ wird mit der Unbekannten $p_{i|I}$ bezeichnet.

Dabei soll jeder gerichtete Zyklus C in der ungerichteten Version von $G(\mathcal{E})$ ein Binom repräsentieren: Jede Kantenbenennung erscheint als Faktor im positiven Teil des Binoms, falls seine Kante in Zyklusrichtung zeigt und im negativem Teil, falls entgegengesetzt.

Beispiel 3.1

- (1) Sei $n = 4$ und $\mathcal{E} = \{12, 123, 1234\}$. Der Graph $G(\mathcal{E})$ ist in Abbildung 3 gezeichnet. Betrachtet man den Zyklus $(1234, 3, 123, 1, 1234)$, so erhält man das Binom $p_3 p_{1|123} - p_{3|123} p_1$ (bzw. $p_3 p_{123} p_1 - p_3 p_{1|123}$).
- (2) Sei $n = 3$ und $\mathcal{E} = \{12, 13, 23, 123\}$. Hier erhält man aus dem äußeren Zyklus (siehe Abbildung 4) das Binom $p_{1|12} p_{3|13} p_{2|23} - p_{2|12} p_{3|23} p_{1|13}$.

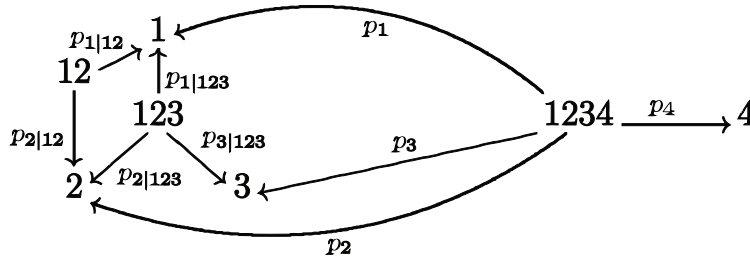


Abbildung 3: Bipartiter Graph $G(\mathcal{E})$ zu $\mathcal{E} = \{12, 123, 1234\}$

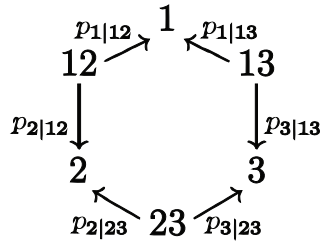


Abbildung 4: Äußerer Zyklus des Graphen $G(\mathcal{E})$ zu $\mathcal{E} = \{12, 13, 23, 123\}$

Ein Zyklus C heißt *induziert*, falls kein Paar seiner Ecken im ursprünglichen Graphen durch eine Kante verbunden ist, die nicht zu C gehört. (“Man kann keine Abkürzung nehmen.”)

Satz 3.2 Die durch den Graphen $G(\mathcal{E})$ definierten Binome bilden eine universelle Gröbnerbasis von $I_{\mathcal{E}}$. Insbesondere wird $I_{\mathcal{E}}$ schon von den Binomen aus induzierten Zyklen erzeugt, wobei diese nicht notwendig eine Gröbnerbasis bilden.

Um diesen Satz zu zeigen, sind einige Aussagen über torische unimodulare Ideale sowie unimodulare Matrizen nötig, die von Sturmfels in [4] behandelt werden. Dazu:

Definition 3.3 Die Matrix $A \in \mathbb{Z}^{m \times d}$ definiert das (sogenannte) torische Ideal

$$I_A := \langle x^{u^+} - x^{u^-} : u \in \text{Bild}(A) \cap \mathbb{Z}^m \rangle,$$

wobei $u = x^{u^+} - x^{u^-}$ mit $x^{u^+}, x^{u^-} \in \mathbb{N}_0^m$.

Definition 3.4 Ein torisches Ideal I_A heißt unimodular, falls jede reduzierte Gröbnerbasis von I_A ausschließlich aus quadradfreien Binomen besteht (d.h. es existieren keine Quadrate in den Leittermen jeder Gröbnerbasis). In diesem Fall heißt auch die Matrix A unimodular.

Ist $G = (U, V, E)$ ein bipartiter Graph, so erhält man ihre Ecken-Kanten-Inzidenzmatrix auf folgende Weise [3]: die Zeilen werden mit den Ecken des Graphen, die Spalten mit den Kanten des Graphen bezeichnet. Ein Eintrag a_{ij} der Matrix soll dann 1 sein, wenn die Ecke i in der Kante j liegt und sonst 0.

Beispiel 3.5 Man betrachte den Graphen G (siehe Abbildung 5) mit

$$V = \{1, 2, 3, 4\}, \quad U = \{\{1, 2\}, \{1, 3\}, \{2, 4\}, \{1, 4\}\}.$$

Dann ist

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

die Ecken-Kanten-Inzidenzmatrix.

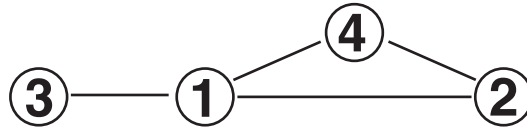


Abbildung 5: Die Ecken-Kanten-Inzidenzmatrix des Graphen G aus Beispiel 3.5

In unserem Fall ist $G = G(\mathcal{E})$ und es gilt:

$$U = \{u_I : I \in \mathcal{E}\} \quad \text{sowie} \quad V = \{v_i : i \in \cup_{I \in \mathcal{E}} I\}.$$

Die Zeilen der Ecken-Kanten-Inzidenzmatrix A des Graphen $G(\mathcal{E})$ werden dann mit den Ecken $u_1, \dots, u_{|U|}, v_1, \dots, v_{|V|}$ bezeichnet, die Spalten wie gehabt.

Die folgende Proposition besagt, dass so konstruierte Matrizen unimodular sind und erlaubt so den Beweis von Satz 3.2.

Proposition 3.6 Die Ecken-Kanten-Inzidenzmatrix A eines bipartiten Graphen $G = (U, V, E)$ ist unimodular, d.h. insbesondere ist I_A ein unimodulares torisches Ideal. Die Zyklus-Binomi des Graphen G bilden eine universelle Gröbnerbasis von I_A .

Man zeigt nun, dass die dem Graphen $G(\mathcal{E})$ assoziierten Binome eine Basis des torischen Ideals I_A bilden, wobei A die Ecken-Kanten-Inzidenzmatrix von $G(\mathcal{E})$ ist. Außerdem zeigt man, dass dieses Ideal mit $I_{\mathcal{E}}$ übereinstimmt. Diesen Beweis und weitergehende Betrachtungen sind in [2] ausgeführt, welches Vorlage für diese Ausarbeitung war.

4 Quellen

- [1] K. HULEK: *Elementare Algebraische Geometrie*. Vieweg, 2000.
- [2] J. MORTON: *Relations among conditional probabilities*. <http://arxiv.org/abs/0808.1149>, 8. August 2008.
- [3] A. SLAVKOVIC und S. SULLIVANT: *The space of compatible full conditionals is a unimodular toric variety*. *Journal of Symbolic Computation*, 41:196-209, 2006.
- [4] B. STURMFELS: *Gröbner Bases and Convex Polytopes*. American Mathematical Society, Providence, 1996.