

Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Piotr Zwiernik

Nr albumu: 201240

Geometria toryczna statystycznych binarnych modeli grafów

Praca magisterska
na kierunku MATEMATYKA
w zakresie MATEMATYKI OGÓLNEJ

Praca wykonana pod kierunkiem
prof. dr. hab. Jarosława Wiśniewskiego
Instytut Matematyki

Grudzień 2006

Oświadczenie kierującego pracą

Potwierdzam, że niniejsza praca została przygotowana pod moim kierunkiem i kwalifikuje się do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data

Podpis kierującego pracą

Oświadczenie autora (autorów) pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data

Podpis autora (autorów) pracy

Streszczenie

Przy zachowaniu stosunkowo dużej ogólności rozważań, możemy patrzeć na modele statystyczne zmiennych dyskretnych, jako na obrazy przekształceń wielomianowych. W ten sposób można wspomóc analizę statystyczną narzędziami geometrii algebraicznej i algebry przemiennej. Badanie parametrycznych algebraicznych modeli statystycznych stało się jednym z ważnych zadań statystyki algebraicznej. Modele graficzne i modele grafów stanowią ważne podrodziny tych modeli, niezwykle często stosowane przez statystyków. Rozmaitości algebraiczne, które pojawiają się w tym kontekście są toryczne (w sensie Sturmfelsa [Stu96]). W ogólności jednak nie muszą być normalne, a zatem nie możemy stosować klasycznej teorii rozmaitości torycznych. Jak pokazali Buczyńska i Wiśniewski [BW06] udowodnienie normalności rozmaitości dla danej rodziny modeli może w znacznym stopniu ułatwić jej badanie.

W pracy tej koncentrujemy się na geometrii pewnej rodziny modeli statystycznych. Po przedstawieniu odpowiedniej części teorii z geometrii torycznej oraz statystyki algebraicznej w rozdziale 3 przedstawimy binarne modele grafów (zmiennie losowe w analizowanym układzie są binarne). Głównym wynikiem pracy jest wykazanie, że dla modelu grafu bez cykli, odpowiadająca mu rozmaitość toryczna jest normalna. Zbadamy również dokładniej przypadek grafów cyklicznych. Jest to próba przybliżenia się do hipotezy postawionej ostatnio przez Sturmfelsa i Sullivanta [SS06], że modelowi graficznemu, który nie zawiera podgrafu pełnego o czterech wierzchołkach, odpowiada rozmaitość normalna.

Słowa kluczowe

geometria toryczna, statystyka algebraiczna, modele graficzne, pokrycie unimodularne, rozmaitość normalna

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.1 Matematyka

11.2 Statystyka

Klasyfikacja tematyczna

Mathematics Subject Classification: 14M25,14Q99.

Tytuł pracy w języku angielskim

Toric geometry of statistical binary graph models

Spis treści

Wprowadzenie	5
1. Geometria algebraiczna	9
1.1. Ideały toryczne i rozmaitości semitoryczne	9
1.2. Wielościany i rozmaitości toryczne	12
1.2.1. Konstrukcja klasyczna	12
1.2.2. Konstrukcja z wielościanu	12
1.2.3. Normalność rozmaitości semitorycznych	14
1.3. Kombinatoryka rozmaitości torycznych z wielościanów	16
1.4. Przykład podsumowujący	17
2. Statystyka algebraiczna	21
2.1. Model statystyczny	21
2.2. Geometria modeli statystycznych	23
2.3. Modele toryczne i modele niezależności	25
2.4. Modele graficzne i modele grafów	28
3. Geometria binarnych modeli grafów	33
3.1. Model i parametryzacja	33
3.2. Normalność wielościanów niektórych modeli grafów	37
3.2.1. Unimodularność dla modeli grafów liniowych	37
3.2.2. Unimodularność dla modeli grafów bez cykli	40
3.2.3. Kwestia unimodularności dla modeli cyklicznych	43
3.3. Więcej o wielościanie P_n	45
3.4. Rozmaitość Y_n modelu \mathcal{C}_n dla małych n	51
4. Zakończenie	55
A. Lematy o unimodularności	57
B. Obliczenia	63
B.1. Wielościany P_3 i R_3	63
B.2. Obliczenia w Polymake	64
B.2.1. Wielościan P_3	64
B.2.2. Wielościan P_4	65
B.3. Użyteczny skrypt dla Polymake	66

Wprowadzenie

Statystyka algebraiczna to dział matematyki, który zajmuje się wykorzystaniem narzędzi algebry przemiennej i geometrii algebraicznej w analizie statystycznej. Podstawowa idea polega na spojrzeniu na model statystyczny jako na zbiór algebraiczny. Techniki geometro-algebraiczne dają szczególnie wiele w przypadku skończonych przestrzeni stanów. Dodatkowo w przypadku dyskretnym uzyskujemy bardzo dużą ogólność rozważań, gdyż przez prostą zamianę zmiennych możemy patrzeć na przykład na dowolną rodzinę wykładniczą jako na przekształcenie wielomianowe.

Sama nazwa statystyki algebraicznej pojawiła się po raz pierwszy w monografii Giovanni Pistone, Eva Riccomagno i Henry Wynn [PRW01]. Książka ta prezentuje głównie podejście algebraiczne. Jej główną zasługą jest przetłumaczenie wielu podstawowych modeli i obiektów statystyki na język algebry. Powszechnie znane narzędzia algebry przemiennej, jak na przykład bazy Gröbnera, mogą być wówczas wykorzystane do efektywnych obliczeń w statystyce.

Równie ważny w rozwoju tej dziedziny okazał się artykuł Persi Diaconisa i Berndta Sturmfelsa [DS98], w którym autorzy zaproponowali algorytm generowania próbek losowych ze stosunkowo szerokiej klasy dyskretnych rozkładów warunkowych pod warunkiem statystyki dostatecznej¹. W ich rozważaniach okazało się, że znalezienie odpowiedniego nieprzywiedlnego łańcucha Markowa na zadanej przestrzeni, odpowiada znalezieniu bazy generatorów dla pewnego ideału torycznego. Podejście to nie tylko pozwoliło efektywniej dokonywać obliczeń, co zapewniło możliwość rozwiązania problemu, który dotąd w ogóle nie był rozwiązany. Było to zatem historycznie pierwsze tak spektakularne osiągnięcie statystyki algebraicznej.

Idealy toryczne stanowią pośrednio jeden z ważnych obiektów zainteresowań tej pracy. Jak się bowiem okazało, pojawiają się one w sposób naturalny w wielu rozważaniach objętych zainteresowaniem statystyki algebraicznej. Jest tak również w przypadku analizowanym w tej pracy. Ciekawym przeglądem najważniejszych obszarów, w których analiza algebry modeli statystycznych prowadzi do badania odpowiednich ideałów torycznych jest praca doktorska obroniona ostatnio przez Seta Sullivanta [Sul05].

Statystyka algebraiczna rozwija się bardzo szybko. Coraz większą rolę gra w tej teorii geometria algebraiczna. Pojawiły się nowe ciekawe powiązania z geometrią toryczną, rozmaitościami stycznymi czy geometrią tropikalną. W obszarze zastosowań szczególną rolę zajmuje projektowanie eksperymentów oraz obliczeniowa biologia molekularna. Przegląd najnowszych metod statystyki algebraicznej w biostatystyce można znaleźć w niedawno wydanej książce [SP05].

Tempo znajdowania kolejnych zastosowań dla algebry i geometrii na polu statystyki każe wierzyć, że jest to dopiero początek tego egzotycznego mariażu. Wydaje się, że statystyka może jeszcze bardzo wiele skorzystać dzięki metodom geometro-algebraicznym. Choć nie w sposób tak efektowny, ale następuje również przepływ idei w drugą stronę - to znaczy od staty-

¹Najbardziej popularne zastosowanie obejmuje znajdowanie p-value testów dla modeli hierarchicznych tabel kontyngencji. Statystyką dostateczną wówczas jest zestaw odpowiednich sum brzegowych tablicy (patrz np. Agresti [Agr02]).

styki i jej zastosowań, do czystej matematyki. Analiza statystycznych modeli algebraicznych, o których będzie również mowa w tej pracy, prowadzi do ciekawych obiektów geometrycznych lub uogólnień tych klasycznych, a także do nowych interesujących twierdzeń. Przegląd takich przypadków z punktu widzenia kombinatoryka i algebraika zaprezentował Sturmfels [Stu06].

Praca ta została podzielona na trzy części. Na początku przedstawimy rozszerzenie klasycznej teorii geometrii algebraicznej, które wydaje się lepiej dopasowane do analizy zaprezentowanej w dalszych rozdziałach. Rozważania te będą oparte na monografii Sturmfelsa [Stu96]. Skupimy się przede wszystkim na kombinatorycznych własnościach rozmaitości torycznych powstających z wielościanów kratowych. Jak się okaże, w tym punkcie część teorii związanej z klasyczną geometrią toryczną (patrz np. Fulton [Ful93]) styka się z podejściem zaprezentowanym przez Sturmfelsa, w którym za cenę bardziej przejrzystego zapisu kombinatorycznego zrezygnowano z założenia normalności rozmaitości torycznych².

W rozdziale 2 przedstawimy podstawowe idee i definicje statystyki oraz statystyki algebraicznej. Ta część pracy ma posłużyć głównie jako motywacja badań zaprezentowanych w ostatnim rozdziale. Ma ona również mieć poniekąd charakter wprowadzający i przeglądowy, gdyż według naszej wiedzy, w polskiej literaturze brak opracowań dotyczących tej tematyki. Wpierw przedstawimy preliminaria statystyczne, wprowadzając pewne podstawowe pojęcia, których będziemy używać w dalszej części rozdziału. Następnie w podrozdziale 2.2 zaprezentujemy geometryczne podejście do modeli statystycznych. W podrozdziale 2.3 przedstawimy modele toryczne, które odpowiadają statystycznej rodzinie modeli wykładniczych, a które są w bezpośredni sposób związane z zagadnieniami geometrii torycznej. Wykorzystując rozważania rozdziału 1 pokażemy, jak przekształcenie wielomianowe modelu torycznego definiuje rozmaitość semitoryczną. Następnie w podrozdziale 2.4 przedstawimy modele graficzne, oraz modele grafów. Są to dwie bardzo ważne i popularne klasy modeli statystycznych, którym poświęcono już kilka monografii (patrz np. Whittaker [Whi90], Lauritzen [Lau96] i Edwards [Edw00]).

Głównym celem tej pracy jest możliwie dogłębna analiza pewnych podklas modeli grafów, a mianowicie binarnych modeli grafów cyklicznych oraz binarnych modeli grafów bez cykli. Te pierwsze będziemy skrótowo nazywać *modelami* \mathcal{C}_n , gdzie n oznacza liczbę wierzchołków w grafie cyklicznym. Od strony kombinatorycznej dużo pracy wykonali Sullivant i Hoşten [HS02]. My w rozdziale trzecim skoncentrujemy się na aspekcie geometrycznym. Podstawowym wynikiem tego rozdziału jest (wniosek 3.11) udowodnienie, że dla dowolnego grafu bez cykli, modelowi tego grafu odpowiada normalna rozmaitość semitoryczna, a zatem toryczna. Wynik ten bazuje między innymi na lemacie o unimodularności dla wielościanu kratowego będącego produktem włóknistym nad simpleksem dwóch normalnych wielościanów kratowych. Lemat ten jest uogólnieniem wyniku Buczyńskiej i Wiśniewskiego [BW06] i został zaprezentowany w dodatku A. Na mocy hipotezy wysuniętej ostatnio przez Sturmfelsa i Sullivanta [SS06], że modelowi graficznemu, który nie zawiera podgrafu pełnego o czterech wierzchołkach, odpowiada normalna rozmaitość toryczna, możemy oczekiwać, że również modele grafów cyklicznych zadają rozmaitość normalną. Obliczenia potwierdzają tę hipotezę dla niskiej liczby wierzchołków grafu. Wierząc, że w dalszych badaniach ważne będzie zrozumienie, jak wygląda wielościan zadany przez binarny model cykliczny, zaprezentujemy serię wyników, które stanowią weryfikację oraz pogłębienie wyników zaprezentowanych przez Sullivanta i Hoštena w artykule [HS02]. Na koniec przedstawimy podstawowe charakterystyki niskowymiarowych rozmaitości modelu \mathcal{C}_n . W obliczeniach tego podrozdziału w niezwykle intensywny

²Mylące może być nazywanie jednego obiektu rozmaitością toryczną, jeśli według innej teorii rozmaitością toryczną nie jest. Dlatego rozmaitość definiowaną przez Sturmfelsa będziemy nazywać rozmaitością semitoryczną, tzn. bez założenia normalności.

sposób korzystaliśmy z programu `Polymake` [GJ05]. Dzięki autorom tego programu udało się nam również uzyskać użyteczny skrypt, który pozwala zautomatyzować liczenie objętości simpleksów z triangulacji wielościanu kratowego (patrz dodatek B.3)

W pracy tej staraliśmy się możliwie dokładnie przedstawić podstawy niezbędne do przedstawienia właściwej części opisywanej teorii. Ograniczona forma wymusza jednak zrezygnowanie z opisu niektórych wykorzystywanych zagadnień. Będziemy na przykład zakładać znajomość podstaw teorii grafów (nieskierowanych), geometrii algebraicznej oraz geometrii dyskretnej.

Rozdział 1

Geometria algebraiczna

W tej części pracy na początek opisujemy krótko rozmaitości algebraiczne, które są parametryzowane przez jednomiany. Rozmaitości tego typu będziemy nazywać rozmaitościami semitorycznymi. Następnie zaprezentujemy użyteczną część klasycznej teorii rozmaitości torycznych, w duchu książki Fultona [Ful93]. Na koniec skoncentrujemy się na geometrii i kombinatoryce rozmaitości torycznych konstruowanych z wachlarza normalnego wielościanu kratowego. Rozważania te posłużą nam w dalszej części tekstu do zdefiniowania rozmaitości torycznej oraz semitorycznej dla statystycznego binarnego modelu grafu cyklicznego.

1.1. Ideały toryczne i rozmaitości semitoryczne

W tej pracy będziemy interesować się specjalną klasą ideałów w pierścieniu $\mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, \dots, x_n]$, które będziemy nazywać ideałami torycznymi. Ideały te definiują rozmaitości, które będziemy nazywać rozmaitościami semitorycznymi. Ta część teorii przedstawiona jest szczegółowo w książce Sturmfelsa [Stu96]. Rozmaitości semitoryczne stanowią uogólnienie rozmaitości torycznych. W tym podrozdziale zaprezentujemy sposób ich konstrukcji.

Niech A będzie macierzą wymiaru $d \times n$ o współczynnikach w zbiorze liczb całkowitych. Każdą kolumnę a_i ($i = 1, \dots, n$) tej macierzy będziemy identyfikować z odpowiadającym jej jednomianem Laurenta w pierścieniu $\mathbb{C}[t_1, t_1^{-1}, \dots, t_d, t_d^{-1}]$ w ten sposób, że $a_i \sim \mathbf{t}^{a_i} = t_1^{a_{1i}} \cdots t_d^{a_{di}}$. Ideałem torycznym \mathcal{I}_A związanym z A będziemy nazywać jądro homomorfizmu \mathbb{C} -algebr

$$\psi_A : \mathbb{C}[x_1, x_2, \dots, x_n] \rightarrow \mathbb{C}[t_1, t_1^{-1}, \dots, t_d, t_d^{-1}], \quad x_i \mapsto \mathbf{t}^{a_i}. \quad (1.1)$$

Jeżeli potraktujemy macierz A jako macierz przekształcenia \mathbb{Z} -modułów, to okazuje się, że \mathcal{I}_A ma przedstawienie w terminach wektorów z jądra tego przekształcenia. Dla każdego $\mathbf{u} \in \mathbb{Z}^n$ możemy zapisać $u = u^+ - u^-$, gdzie u^+ i u^- mają nieujemne współrzędne i rozłączne nośniki. Każde takie wyrażenie w dobrze zadaje nam różnicę dwóch jednomianów, którą będziemy nazywać dwumianem.

Lemat 1.1 ([Stu96], **lemat 4.1**) *Ideal toryczny \mathcal{I}_A jest generowany przez dwumiany $x^{u^+} - x^{u^-}$, gdzie u przebiega $\ker_{\mathbb{Z}}(A)$.*

Dowód. Niech $I = \langle x^{u^+} - x^{u^-} \mid u \in \ker_{\mathbb{Z}}(A) \rangle$. Oczywiście $I \subset \mathcal{I}_A$, bo skoro $u \in \ker_{\mathbb{Z}}(A)$ to $Au^+ = Au^-$ oraz, jak łatwo sprawdzić $\psi_A(x^v) = x^{Av}$ dla każdego $v \in \mathbb{Z}_{\geq 0}^n$. Zatem $x^{u^+} - x^{u^-} \in \ker(\psi_A)$

Ustalmy porządek na jednomianach taki, że zmienne porządkujemy liniowo względem indeksu, to znaczy $x_1 > \dots > x_n$. Dodatkowo pierwszy jednomian jest większy od drugiego

jeśli albo stopień pierwszego jest większy albo gdy stopnie są równe stosujemy porządek leksykograficzny.

Przypuśćmy, że inkluzja $I \subset \mathcal{I}_A$ jest ścisła. Niech $f \in \mathcal{I}_A \setminus I$ oraz niech x^α będzie największym jednomianem f (w sensie zadanego porządku). Skoro $\psi_A(f) = 0$ musi istnieć drugi jednomian x^β w f taki, że $\psi_A(x^\alpha) = \psi_A(x^\beta)$. Możemy wydzielić wspólne czynniki tych jednomianów i dostajemy $x^\alpha - x^\beta = x^\gamma(x^{\alpha'} - x^{\beta'})$ oraz α' i β' mają rozłączne nośniki. Oczywiście $\psi_A(x^{\alpha'}) - \psi_A(x^{\beta'}) = 0$. Niech $h = \alpha' - \beta'$ i oczywiście $h \in \ker(A)$ a zatem $x^\gamma(x^{h^+} - x^{h^-}) \in I$, gdzie $h^+ = \alpha'$ oraz $h^- = \beta'$, bo drugi z tych czynników leży w I . Odejmując ten wyraz od f , dostajemy wielomian w $\mathcal{I}_A \setminus I$, którego największy jednomian jest mniejszy niż x^α . Procedurę możemy powtarzać pokazując tym samym, że ideały I i \mathcal{I}_A w istocie muszą być równe. \square

Oczywiście ideał \mathcal{I}_A jest ideałem pierwszym jako jądro homomorfizmu idącego w dziedzinę całkowitości, a zatem zbiór jego zer w \mathbb{C}^n definiuje nieredukowalną rozmaitość afiniczną (patrz [Har77], rozdział I, wniosek 1.4). Zbiór ten będziemy oznaczać jako X_A i nazywać afiniczną rozmaitością semitoryczną, zadaną przez A . Wprost z definicji wynika, że X_A jest domknięciem Zariskiego obrazu przekształcenia

$$\phi: (\mathbb{C}^*)^d \rightarrow \mathbb{C}^n, \quad (t_1, \dots, t_d) \mapsto (\mathbf{t}^{a_1}, \dots, \mathbf{t}^{a_n}). \quad (1.2)$$

Widać zatem również, że X_A zawiera torus (obraz $(\mathbb{C}^*)^d$ względem powyższego przekształcenia), który oczywiście jest gęsty w X_A . Torus $(\mathbb{C}^*)^d$ działa na \mathbb{C}^n poprzez

$$(x_1, \dots, x_n) \mapsto (x_1 \mathbf{t}^{a_1}, \dots, x_n \mathbf{t}^{a_n}).$$

Zatem inaczej afiniczna rozmaitość semitoryczna X_A to domknięcie $(\mathbb{C}^*)^d$ -orbity punktu $(1, \dots, 1)$.

Za Sturmfelsem [Stu96] podajemy sposób na wyliczenie wymiaru rozmaitości X_A . Kratę generowaną przez kolumny macierzy A będziemy oznaczać przez M . Wymiar kraty M równy jest $\text{rz}(A)$.

Lemat 1.2 ([Stu96], lemat 4.2) *Wymiar Krulla pierścienia $k[\mathbf{x}]/\mathcal{I}_A$ równy jest $\text{rz}(A)$.*

Dowód: Pierścień $k[\mathbf{x}]/\mathcal{I}_A$ jest izomorficzny z podpierścieniem $k[\mathbf{t}^{a_1}, \dots, \mathbf{t}^{a_n}]$ pierścienia $k[\mathbf{t}^{\pm 1}]$. Wymiar Krulla tej dziedziny jest równy maksymalnej liczbie algebraicznie niezależnych jednomianów \mathbf{t}^{a_i} . Ale zbiór jednomianów jest algebraicznie niezależny wtedy i tylko wtedy, gdy odpowiadający im zbiór wektorów w wykładnikach jest liniowo niezależny. \square

Powyższy lemat oraz znany wynik (patrz [Har77], rozdział I, twierdzenie 1.7) zapewniają, że wymiar rozmaitości X_A jest równy rzędowi macierzy A .

Głównym obszarem naszego zainteresowania w tej pracy będą semitoryczne rozmaitości rzutowe. Ich konstrukcja przebiega analogicznie do przedstawionej powyżej. Niech $A = (a_1, \dots, a_n)$ będzie jak wcześniej. Semitoryczną rozmaitością rzutową $Y_A \subset \mathbb{P}^{n-1}$ będziemy nazywać rozmaitość związaną z ideałem generowanym przez wszystkie jednorodnie dwumiany ideału $\mathcal{I}_A \subset \mathbb{C}[x_1, \dots, x_n]$. Najbardziej interesujący przypadek – również z punktu widzenia dalszej części pracy – otrzymujemy, gdy przestrzeń rozpinana przez wiersze macierzy A zawiera wektor jedynek $[1, \dots, 1]$, to znaczy, gdy leży on w obrazie przekształcenia A^T traktowanego jako przekształcenie odpowiednich przestrzeni liniowych. W tej sytuacji bowiem dwumiany spełniające warunek z lematu 1.1 są automatycznie jednorodnie, a zatem ideał toryczny \mathcal{I}_A jest ideałem jednorodnym rozmaitości Y_A . Wynika to z tego, że $\ker(A)^\perp = \text{im}(A^T)$, a zatem jeśli $u \in \ker(A)$ to $u \cdot [1, \dots, 1] = 0$, bo $[1, \dots, 1] \in \text{im}(A^T)$, a zatem $|u^+| = |u^-|$, gdzie $|u|$ to suma współrzędnych wektora u . Zatem $x^{u^+} - x^{u^-}$ jest jednorodny.

Przykład 1.3 (Kubika w \mathbb{P}^3) *Niech*

$$A = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

Ideal toryczny \mathcal{I}_A generowany jest przez trzy dwumiany kwadratowe:

$$\mathcal{I}_A = \langle x_0x_2 - x_1^2, x_0x_3 - x_1x_2, x_1x_3 - x_2^2 \rangle.$$

Zera tego ideału zadają kubikę w \mathbb{P}^3 .

Przypadek rzutowy wiąże się z tak zwanymi wielościennymi pierścieniami grupowymi. Jest to ważne pojęcie, gdyż wiele ciekawych własności rzutowych rozmaitości semitorycznych daje się odczytać z analizy tych pierścieni. Konstrukcja wielościennego pierścienia półgrupowego przebiega następująco. Niech P będzie wielościannem kratowym w \mathbb{Z}^d i niech dane będzie zanurzenie $\iota : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$, $\iota(x) = (x, 1)$. Definiujemy S_P jako półgrupę generowaną przez punkty kratowe w $\iota(P)$ i będziemy nazywać półgrupą wielościenną. Dla dowolnej półgrupy Q definiujemy pierścień półgrupowy $\mathbb{C}[Q]$, jako \mathbb{C} -algebrę z \mathbb{C} -bazą złożoną z elementów formalnych $\{\mathbf{t}^a : a \in Q\}$ oraz mnożeniem zdefiniowanym tak, że $\mathbf{t}^a \cdot \mathbf{t}^b = \mathbf{t}^{a+b}$. Z definicji \mathbb{C} -algebrę $\mathbb{C}[S_P]$ będziemy nazywać wielościennym pierścieniem półgrupowym (patrz [BGT97]). Poniższe twierdzenie mówi o związku wielościennych pierścieni półgrupowych z teorią rozmaitości semitorycznych.

Twierdzenie 1.4 ([BGT97]) *Niech R będzie dziedziną całkowitości. R jest wielościennym pierścieniem półgrupowym (jest izomorficzny z wielościennym pierścieniem półgrupowym) wtedy i tylko wtedy, gdy R jest pierścieniem z gradacją $R = \bigoplus_{i=0}^{\infty} R_i$ takim, że*

(i) $R_0 = \mathbb{C}$ oraz R jest \mathbb{C} -algebrą skończenie generowaną przez $a_1, \dots, a_n \in R_1$,

(ii) jądro naturalnego epimorfizmu

$$\varphi : \mathbb{C}[x_1, \dots, x_n] \rightarrow R, \quad \varphi(x_i) = a_i$$

jest generowane przez dwumiany $x^u - x^v$, gdzie $x^u = x_1^{u_1} \cdots x_n^{u_n}$ dla $u = (u_1, \dots, u_n) \in \mathbb{Z}_{\geq 0}^n$,

(iii) $R_1 = \bar{R}_1$, gdzie przez \bar{R} oznaczamy normalizację pierścienia R .

Powyższe twierdzenie pokazuje, że wielościenny pierścień półgrupowy wiąże się bezpośrednio z semitorycznymi rozmaitościami rzutowymi. Rzeczywiście, jeśli punkty kratowe wielościannu $\iota(P)$ oznaczymy jako $a_1, \dots, a_n \in \mathbb{Z}^d \oplus \{1\}$ oraz generatory algebry $\mathbb{C}[S_P]$ formalnie jako $\mathbf{t}^{a_j} = \prod_{i=1}^{d+1} t_d^{a_{ij}}$ dla $j = 1, \dots, n$, to na mocy punktu (ii) twierdzenia dostajemy, że jądro φ jest ideałem torycznym, a więc definiuje rozmaitość semitoryczną. Jeśli współrzędne wierzchołków $\iota(P)$ zapiszemy w kolejnych kolumnach macierzy A , to oczywiście wektor $[1, \dots, 1]$ leży w przestrzeni $\text{im}(A^T)$, a zatem skonstruowany powyżej ideał \mathcal{I}_A jest jednorodny na mocy rozważań z podrozdziału 1.1. A zatem definiowana rozmaitość jest rzutowa. Ważne, że rozumowanie da się odwrócić. Jeśli mamy wielościann P , którego macierz współrzędnych A spełnia $[1, \dots, 1] \in \text{im}(A^T)$ to na mocy konstrukcji przedstawionej powyżej definiuje on wielościenny pierścień półgrupowy $\mathbb{C}[S_P]$ i w konsekwencji semitoryczną rozmaitość rzutową, co wynika z twierdzenia 1.4.

1.2. Wielościany i rozmaitości toryczne

W podrozdziale 1.1 używaliśmy pojęcia rozmaitości semitorycznej na określenie zbioru zer dowolnego ideału torycznego w pierścieniu $\mathbb{C}[x_1, \dots, x_n]$. Takie podejście jednak nie jest zgodne ze standardową definicją rozmaitości torycznej (patrz np. Fulton [Ful93]) w przestrzeni \mathbb{C}^n . Klasyczna definicja brzmi następująco.

Definicja 1.5 ([KKMSD73]) *Rozmaitością toryczną nazywamy nieredukowalną rozmaitość X taką, że*

1. X jest normalna
2. $(\mathbb{C}^*)^n$ jest otwartym (w topologii Zariskiego) podzbiorem X , oraz
3. działanie torusa $(\mathbb{C}^*)^n$ na samym sobie rozszerza się do działania $(\mathbb{C}^*)^n$ na X .

Również standardowa konstrukcja rozmaitości torycznej różni się od tej przedstawionej w poprzednim podrozdziale. Teraz przedstawimy, jak połączyć oba podejścia. Będziemy korzystać z notacji przedstawionej w książce Fultona [Ful93].

1.2.1. Konstrukcja klasyczna

Ustalmy kratę $N \simeq \mathbb{Z}^d$ i kratę dualną $M = N^\vee = \text{Hom}(N, \mathbb{Z})$ oraz ich uliniowania $N_{\mathbb{R}} = N \otimes_{\mathbb{Z}} \mathbb{R}$ i $M_{\mathbb{R}} = M \otimes_{\mathbb{Z}} \mathbb{R}$. Niech σ oznacza ściśle wypukły wymierny stożek wielościany w przestrzeni wektorowej $N_{\mathbb{R}} \simeq \mathbb{R}^d$, czyli stożek o początku w układzie współrzędnych, rozpięty przez skończoną liczbę wektorów kratowych (tzn. $\sigma = \sum_{i=1}^k \mathbb{R}_{\geq 0} v_i$) i taki nie zawierający linii przechodzącej przez środek układu współrzędnych. Wobec tego stożek dualny do niego

$$\sigma^\vee := \{u \in M_{\mathbb{R}} : \langle u, v \rangle \geq 0 \text{ dla wszystkich } v \in \sigma\}$$

jest d -wymiarowym stożkiem w dualnej przestrzeni wektorowej. Ze stożkiem dualnym będziemy wiązać półgrupę

$$S_\sigma := \sigma^\vee \cap M = \{u \in M : \langle u, v \rangle \geq 0 \text{ dla wszystkich } v \in \sigma\}.$$

Ma ona skończoną liczbę generatorów, a zatem również algebra półgrupowa $\mathbb{C}[S_\sigma]$ jest skończenie generowana (patrz [Ful93], s. 12). Jej spektrum maksymalne $U_\sigma := \text{Spec}(\mathbb{C}[S_\sigma])$ jest normalną afiniczną rozmaitością toryczną. Jeżeli τ jest ścianą σ , to S_σ jest zawarte w S_τ , a zatem $\mathbb{C}[S_\sigma]$ jest podalgebrą w $\mathbb{C}[S_\tau]$, co zadaje przekształcenie $U_\tau \rightarrow U_\sigma$. Rozmaitość U_τ stanowi w istocie otwarty podzbiór U_σ . Istotnie, wybierzmy $u \in S_\sigma$ takie, że $\tau = \sigma \cap u^\perp$, wówczas $U_\tau = \{x \in U_\sigma : u(x) \neq 0\}$. W ogólności pozwala to skonstruować abstrakcyjną rozmaitość toryczną $X(\Sigma)$ z wachlarza Σ w N , który jest zbiorem ściśle wypukłych wymiernych stożków wielościanowych, spełniający dodatkowe własności: (i) każda ściana stożka w Σ jest także stożkiem w Σ , (ii) przecięcie dwóch stożków w Σ jest ich wspólną ścianą. Konstrukcję rozmaitości torycznej $X(\Sigma)$ przeprowadza się poprzez wzięcie sumy rozłącznej rozmaitości afinicznych U_σ dla wszystkich $\sigma \in \Sigma$ oraz sklejanie U_σ i U_τ wzdłuż wspólnych podzbiorów otwartych zdefiniowanych przez stożek $\sigma \cap \tau$.

1.2.2. Konstrukcja z wielościanu

Od tej chwili w naszych rozważaniach będziemy ograniczać się do przypadku, gdy wachlarz Σ w N jest wachlarzem normalnym pewnego wielościanu P w $M_{\mathbb{R}}$ o wierzchołkach w kracie

M . Będziemy zakładać, że $\dim(P) = d$ (wielościan jest pełnowymiarowy), ale nie zakładamy, że P zawiera środek układu współrzędnych. Wachlarz normalny Σ_P zadany jest przez wielościan P w następujący sposób. Dla każdej ściany F wielościanu P mamy stożek σ_F w Σ_P zdefiniowany jako:

$$\sigma_F = \{v \in N_{\mathbb{R}} : \langle u, v \rangle \leq \langle u', v \rangle \text{ dla wszystkich } u \in F \text{ oraz } u' \in P\}$$

Innymi słowy, σ_F jest stożkiem dualnym do zbioru wektorów o początku w F i końcu w P . Stożek σ_F^\vee jest generowany przez wektory $u' - u$, gdzie u i u' przebiegają odpowiednio zbiór wierzchołków ściany F i wielościanu P . Teoria programowania liniowego pozwala patrzeć na σ_F jako na zbiór funkcjonałów liniowych na $M_{\mathbb{R}}$, które po ograniczeniu do P minimalizują się w F .

Twierdzenie 1.6 ([Ful93], s. 26) *Stożki σ_F , gdzie F odpowiada ścianom wielościanu P , tworzą wachlarz Σ_P w $N_{\mathbb{R}}$.*

Tak skonstruowany wachlarz normalny Σ_P definiuje rzutową rozmaitość toryczną X_P , która sklejana z afinicznymi rozmaitościami torycznymi X_σ , gdzie σ przebiega wszystkie stożki Σ_P . Co więcej mamy wynik odwrotny.

Twierdzenie 1.7 ([Cox], s. 34) *Rozmaitość toryczna $X(\Sigma)$, gdzie Σ jest wachlarzem w $N_{\mathbb{R}} \simeq \mathbb{R}^n$, jest rzutowa wtedy i tylko wtedy, gdy Σ jest wachlarzem normalnym n -wymiarowego wielościanu kratowego w $M_{\mathbb{R}}$.*

A priori nie istnieje żaden model X_P w przestrzeni rzutowej. Aby uzyskać zanurzenie X_P w przestrzeń rzutową musimy wybrać dywizor szeroki na X_P (porównaj podrozdział II.7 w [Har77]). Rozpatrzmy dywizor D_P zdefiniowany w podrozdziale 1.3. Niech P_{D_P} będzie wielościanem w $M_{\mathbb{R}}$, który jest związany z D_P ([Ful93], s. 66). Okazuje się, że w tym przypadku zachodzi $P_{D_P} = P$ ([Ful93], s. 72) i otrzymujemy przekształcenie ψ_{D_P} z X_P do rozmaitości rzutowej Y_A , która jest zbiorem zer jednorodnego ideału torycznego \mathcal{I}_A , gdzie (patrz [Stu96], s. 131) kolumny macierzy A są wektorami ze zbioru \mathcal{A} takiego, że

$$\mathcal{A} = (P \cap M) \times \{1\} \subset M \oplus \mathbb{Z} \simeq \mathbb{Z}^{d+1}. \quad (1.3)$$

Rozważania poprzedniego podrozdziału pokazały, że zapewnia to warunek jednorodności ideału \mathcal{I}_A . To, że Y_A jest rzeczywiście obrazem X_P względem przekształcenia ψ_{D_P} , wynika z twierdzenia 1.14 oraz dyskusji, która się pojawia poniżej niego.

W ten sposób częściowo udało nam się pokazać, w którym miejscu konstrukcja z poprzedniego podrozdziału spotyka się z przedstawioną przed chwilą. W konstrukcji Sturmfelsa wychodziliśmy od macierzy A lub równoważnie od przekształcenia jednomianowego, które definiowały rozmaitość semitoryczną. Załóżmy, że przestrzeń rozpinana przez wiersze macierzy A zawiera wektor jedynek i jej rząd jest równy d . Jak pokazaliśmy wcześniej, A zadaje wówczas d -wymiarową afiniczną rozmaitość semitoryczną $X_A \subset \mathbb{C}^n$ oraz $d-1$ -wymiarową rzutową rozmaitość semitoryczną $Y_A \subset \mathbb{P}^{n-1}$ (X_A jest stożkiem nad Y_A). Jeżeli $P = \text{conv}\{a_1, \dots, a_n\}$ będzie wielościanem będącym kombinacją wypukłą kolumn macierzy A , to dostajemy rozmaitość X_P na mocy konstrukcji przedstawionej powyżej. Skoro wiersze macierzy A rozpinają przestrzeń, w której znajduje się wektor jedynek, to wachlarz Σ_P zadaje X_P jako $d-1$ -wymiarową rzutową rozmaitość toryczną (bo $\dim P = d-1$). W ogólności $Y_A \neq X_P$. X_P jest natomiast normalizacją rozmaitości Y_A i jeśli ta druga jest normalna to zachodzi izomorfizm, co pokażemy w podrozdziale 1.3 korzystając głównie z rozważań poniższego paragrafu.

1.2.3. Normalność rozmaitości semitorycznych

W podrozdziale 1.1 pokazaliśmy między innymi, że podrozmaitość posiada torus, który jest w niej gęsty. Działanie torusa na samym sobie rozszerza się na działanie na całej rozmaitości. W związku z tym na mocy definicji 1.5 rozmaitość semitoryczna, która jest normalna, jest toryczna. Dlatego właśnie problem analizy normalności rozmaitości semitorycznych zostanie przedstawiony w tym podrozdziale. Zagadnienie normalności w tym przypadku jest bezpośrednio związane z problemem badania normalności półgrup afinicznych. Półgrupę S nazywamy afiniczną, jeśli jest skończenie generowana i może być zanurzona w \mathbb{Z}^n dla pewnego $n \in \mathbb{N}$. Grupę różnic elementów z S (czyli najmniejszą grupę zawierającą tę półgrupę) będziemy oznaczać $\text{gp}(S)$. Półgrupę afiniczną nazywamy normalną jeśli każdy element $x \in \text{gp}(S)$ taki, że $cx \in S$ dla pewnego $c \in \mathbb{N}$ należy do S . Normalizacją półgrupy S jest półgrupa \bar{S} złożona ze wszystkich $x \in \text{gp}(S)$, dla których istnieje $c \in \mathbb{N}$ takie, że $cx \in S$. Oczywiście \bar{S} jest półgrupą normalną. Kluczowe, że normalność S jest równoważna normalności pierścienia półgrupowego $\mathbb{C}[S]$ (patrz Bruns, Herzog [BH98]¹).

Prawie natychmiastowym wnioskiem z powyższych rozważań jest następujące twierdzenie dotyczące związku między afinicznymi rozmaitościami semitorycznymi a afinicznymi rozmaitościami torycznymi.

Twierdzenie 1.8 ([Stu96], s. 130) *Niech $A = (a_1, \dots, a_n)$ będzie macierzą wymiaru $d \times n$ o współczynnikach w zbiorze liczb całkowitych. Następujące warunki są równoważne:*

1. *Afiniczna rozmaitość semitoryczna X_A jest normalna.*
2. *Afiniczna rozmaitość semitoryczna X_A jest izomorficzna z X_σ dla pewnego stożka σ w $N_{\mathbb{R}}$.*
3. *Dziedzina całkowitości $\mathbb{C}[A] = \mathbb{C}[x]/\mathcal{I}_A$ jest całkowicie domknięta (w swoim ciele ułamków).*
4. *Półgrupa generowana przez kolumny A jest normalna.*

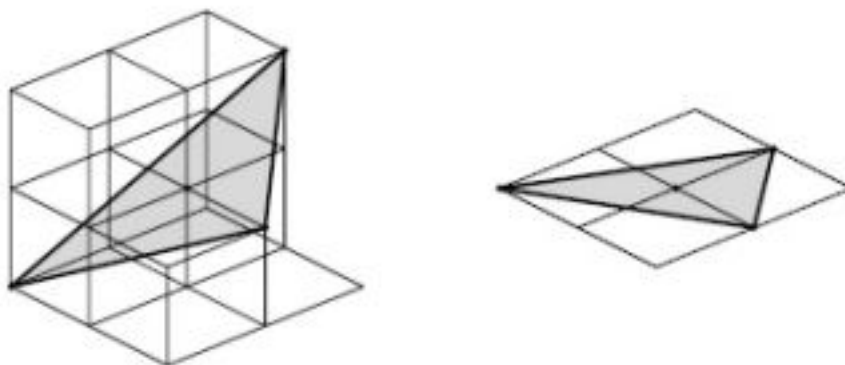
W przypadku rzutowych rozmaitości semitorycznych rozważanie ich normalności związane jest z analizą wielościennych pierścieni półgrupowych zdefiniowanych w poprzednim podrozdziale. Warto zauważyć, że normalność $\mathbb{C}[S_P]$ implikuje normalność rzutową, która jest własnością silniejszą niż normalność i ją implikuje (patrz ćwiczenie 3.18, podrozdział I.3 w [Har77]).

Mówimy, że wielościan kratowy P jest normalny, jeśli półgrupa wielościenna S_P jest normalna.

Twierdzenie 1.9 ([BGT97]) *Niech \mathcal{A} będzie skończonym podzbiorem \mathbb{Z}^n . Niech $E_{\mathcal{A}} = \{(x, 1) : x \in \mathcal{A}\} \subset \mathbb{Z}^{n+1}$. Wówczas przez $C_{\mathcal{A}}$ będziemy oznaczać (wypukły) stożek generowany przez $E_{\mathcal{A}}$, a przez $S_{\mathcal{A}}$ półgrupę generowaną przez $E_{\mathcal{A}}$. Normalizacją \mathbb{C} -algebry $R = \mathbb{C}[S_{\mathcal{A}}]$ jest pierścień półgrupowy $\bar{R} = \mathbb{C}[\text{gp}(S_{\mathcal{A}}) \cap C_{\mathcal{A}}]$. Co więcej względem naturalnej gradacji R i \bar{R} , zachodzi $R_1 = \bar{R}_1$ wtedy i tylko wtedy $\mathcal{A} = P \cap \mathbb{Z}^n$ dla pewnego wielościanu kratowego P .*

Badanie normalności półgrupy wielościennej S_P jest w praktyce trudne. Istnieje jednak stosunkowo prosty do badania warunek implikujący normalność, którym jest istnienie pokrycia unimodularnego dla wielościanu kratowego P .

¹Cytat za [BGT97].



Rysunek 1.1: Dwa kratowo równoważne wielościany (Źródło: [Haa], s. 13)

Definicja 1.10 *Mówimy, że n -wymiarowy simpleks kratowy w \mathbb{R}^n jest unimodularny, jeśli jego objętość wynosi $1/n!$ (znormalizowana objętość wynosi jeden). Równoważnie jeśli simpleks Δ ma wierzchołki v_0, v_1, \dots, v_n w kratce \mathbb{Z}^n to jest unimodularny, jeśli $v_1 - v_0, \dots, v_n - v_0$ tworzą \mathbb{Z} -bazę kraty \mathbb{Z}^n . Zbiór kratowych simpleksów unimodularnych, które pokrywają P nazywamy pokryciem unimodularnym wielościanu P .*

Twierdzenie 1.11 ([BGT97]) *Niech P będzie n wymiarowym wielościanem kratowym w \mathbb{R}^n . Jeżeli P posiada pokrycie unimodularne to jest normalny.*

Jest to bardzo ważne twierdzenie również w kontekście badania rzutowych rozmaitości semitorycznych.

Wniosek 1.12 *Niech P będzie n wymiarowym wielościanem kratowym w \mathbb{R}^n . Jeżeli P posiada pokrycie unimodularne, to ideał toryczny związany z $\mathbb{C}[S_P]$ definiuje rzutowo normalną, a zatem również normalną, semitoryczną rozmaitość rzutową, czyli toryczną rozmaitość rzutową.*

Zauważmy jeszcze założenie w poprzednim twierdzeniu, że wielościan P jest pełnowymiarowy, nie jest szczególnie restrykcyjne. W istocie każdy wielościan jest izomorficzny w sensie izomorfizmu krat z pewnym kratowym wielościanem pełnowymiarowym. Związane jest to z pojęciem kratowej równoważności.

Definicja 1.13 ([Haa]) *Wielościany kratowe P i P' nazywamy kratowo równoważnymi jeśli istnieje przekształcenie afiniczne $\text{aff}(P) \rightarrow \text{aff}(P')$, które przekształca $\mathbb{Z}^d \cap \text{aff}(P)$ bijektywnie na $\mathbb{Z}^d \cap \text{aff}(P')$.*

Istotne jest zauważenie, że dwa kratowo równoważne wielościany definiują tę samą toryczną rozmaitość rzutową.

W tej chwili już przedstawiliśmy odpowiedniość pomiędzy normalnymi afinicznymi rozmaitościami semitorycznymi a afinicznymi rozmaitościami torycznymi. Przedstawiliśmy również sposób badania normalności rzutowych rozmaitości semitorycznych. Pozostaje jedynie pokazać związek między normalnymi rzutowymi rozmaitościami semitorycznymi a rzutowymi rozmaitościami torycznymi. Wynik ten pojawi się pośród rozważań następnego podrozdziału. W ten sposób pokażemy, że istotnie rozmaitości semitoryczne stanowią uogólnienie rozmaitości torycznych.

1.3. Kombinatoryka rozmaitości torycznych z wielościanów

W poprzednim podrozdziale staraliśmy się między innymi pokazać, jak konstrukcja rzutowych rozmaitości semitorycznych łączy się z klasyczną teorią rzutowych rozmaitości torycznych przedstawioną przez Fultona w [Ful93]. Obiektem, który pozwalał połączyć oba podejścia był wielościan w kracie M . W tym podrozdziale zaprezentujemy dalsze własności rozmaitości X_P , które potrafimy odczytać z wielościanu P . Rozważania tego podrozdziału oparte zostały w głównej mierze na czwartym rozdziale notatek Coxa [Cox] i posłużą nam przy analizie geometrii rozmaitości binarnego modelu cyklicznego \mathcal{C}_n w rozdziale 3.

Orbita torusa Pierwszą ważną obserwacją jest, że wielościan P jest kombinatorycznie dualny do swojego wachlarza normalnego Σ_P . To oznacza, że istnieje wzajemnie jednoznaczna odpowiedniość (odwracająca relację zawierania)

$$\sigma_F \in \Sigma_P \longleftrightarrow F \subset P$$

między stożkami w Σ_P a ścianami P (P traktujemy jako właściwą ścianę samego siebie) taka, że:

$$\dim \sigma_F + \dim F = n$$

dla wszystkich ścian $F \subset P$. Na mocy podobnej zależności pomiędzy stożkami Σ_P a orbitami torusa w X_P (patrz [Ful93], s. 51), dostajemy wzajemnie jednoznaczność zachowującą wymiar odpowiedniość pomiędzy ścianami P a orbitami torusa w X_P . W szczególności:

- wierzchołki $P \longleftrightarrow n$ -wymiarowe stożki w $\Sigma_P \longleftrightarrow$ punkty stałe działania torusa w X_P
- ściany P maksymalnego wymiaru \longleftrightarrow promienie w $\Sigma_P \longleftrightarrow$ nieredukowalne dywizory w X_P niezmiennicze względem działania torusa

Dywizory szerokie Jeżeli $P \subset M_{\mathbb{R}} \simeq \mathbb{R}^d$ jest d -wymiarowym wielościanem kratowym to P definiuje wachlarz normalny Σ_P , co opisaliśmy w poprzednim podrozdziale. Jednak P niesie ze sobą więcej informacji o ścianach niż tylko współrzędne wektorów normalnych. Reprezentacja P za pomocą hiperpłaszczyzn podpierających zawiera również liczby całkowite a_F , które występują w każdym z definiujących równań

$$\langle m, n_F \rangle = -a_F$$

hiperpłaszczyzny podpierającej ścianę maksymalnego wymiaru F . Skoro każda taka ściana definiuje nieredukowalny dywizor D_F niezmienniczy względem działania torusa na X_P , to P zadaje dywizor:

$$D_P = \sum_F a_F D_F, \tag{1.4}$$

gdzie suma przebiega po wszystkich ścianach maksymalnego wymiaru P .

Dywizor D_P jest Cartier na X_P (patrz [Cox], tw. 4.2). Dodatkowo D_P jest szeroki. Dokładniej, niech

$$H^0(X_P, \mathcal{O}_{X_P}(D_P)) = \{f \in \mathbb{C}(X_P)^* \mid \operatorname{div}(f) + D_P \geq 0\}.$$

Okazuje się wówczas, że funkcje wymierne pojawiające się w powyższej definicji są jednoznacznie zadane przez punkty kratowe wielościanu P .

Twierdzenie 1.14 ([Cox], tw. 4.3) $H^0(X_P, \mathcal{O}_{X_P}(D_P)) = \bigoplus_{m \in P \cap M} \mathbb{C} \cdot \chi^m$.

Dywizory służą nam do definiowania przekształceń w przestrzeń rzutową. Dla D_P będziemy używać funkcji wymiernych z $H^0(X_P, \mathcal{O}_{X_P}(D_P))$, które na mocy twierdzenia 1.14 są charakterami zadanymi przez punkty kratowe wielościanu P . Dokładniej, niech $N \simeq M \simeq \mathbb{Z}^n$, tak, że punkt kratowy $m \in P \cap \mathbb{Z}^n$ zadaje charakter $\chi^m = t^m$ na torusie $(\mathbb{C}^*)^n$ rozmaitości X_P . Niech m_1, \dots, m_l będą punktami kratowymi wielościanu P . Wówczas dostajemy przekształcenie

$$\phi(t_1, \dots, t_n) = (t^{m_1}, \dots, t^{m_l}) \in \mathbb{P}^{l-1}, \quad l = |P \cap \mathbb{Z}^n|,$$

z $(\mathbb{C}^*)^n$ do \mathbb{P}^{l-1} . Jak widać, jest to szczególny przypadek przekształcenia 1.2 i definiuje rzutową rozmaitość semitoryczną $Y_{P \cap \mathbb{Z}^n}$, o której pisaliśmy w ostatnim akapicie podrozdziału 1.1.

Rozmaitości X_P i $Y_{P \cap \mathbb{Z}^n}$ są ściśle ze sobą związane. Zauważmy, że jeśli $\nu \in \mathbb{Z}_{>0}$ to νP definiuje ten sam wachlarz normalny, a zatem tą samą rozmaitość toryczną, tzn. $X_P = X_{\nu P}$. Łatwo zauważyć również, że $D_{\nu P} = \nu D_P$. Zatem $H^0(X_P, \mathcal{O}_{X_P}(\nu D_P))$ zadaje przekształcenie

$$\phi_\nu : (\mathbb{C}^*)^n \rightarrow \mathbb{P}^{l_\nu-1}, \quad l_\nu = |(\nu P) \cap \mathbb{Z}^n|,$$

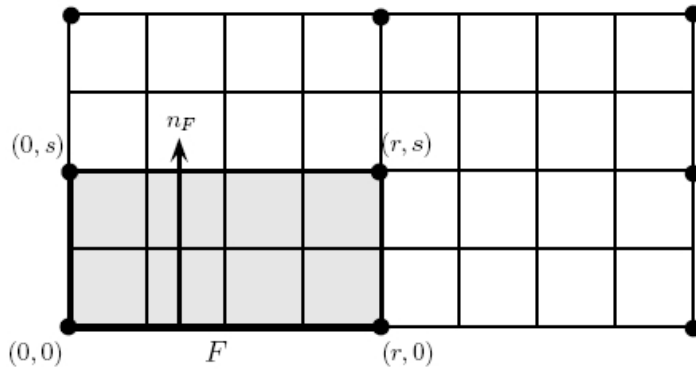
zdefiniowane przez punkty kratowe wielościanu νP . Można pokazać ([Cox], s. 36), że dla $\nu \gg 0$ przekształcenie ϕ_ν indukuje izomorfizm

$$X_P \simeq Y_{(\nu P) \cap \mathbb{Z}^n}.$$

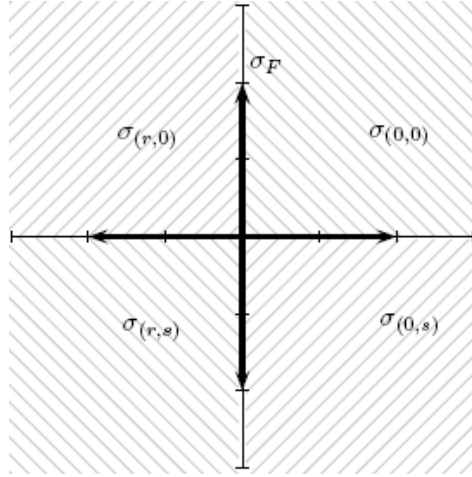
Innymi słowy $H^0(X_P, \mathcal{O}_{X_P}(\nu D_P))$ zadaje zanurzenie X_P w przestrzeń rzutową dla $\nu \gg 0$.

1.4. Przykład podsumowujący

Niech $r, s \in \mathbb{N}$ i niech $P_{r,s} \subset \mathbb{R}^2$ będzie prostokątem kratowym o wierzchołkach $(0, 0)$, $(r, 0)$, $(0, s)$, (r, s) , jak na poniższym rysunku.



Wektory normalne do ścian maksymalnego wymiaru to $n_1 = e_1$, $n_2 = e_2$, $n_3 = -e_1$ oraz $n_4 = -e_2$, gdzie e_i to wektory jednostkowe płaszczyzny \mathbb{R}^2 . Wachlarz normalny $\Sigma_{r,s}$ dla tego wielościanu przedstawiony jest na rysunku poniżej. Cztery stożki maksymalnego wymiaru w wachlarzu normalnym odpowiadają czterem rozmaitościom afinicznym izomorficznymi z \mathbb{C}^2 , dla których odpowiednie definiujące algebry dane są odpowiednio jako $\mathbb{C}[x, y]$, $\mathbb{C}[x^{-1}, y]$, $\mathbb{C}[x, y^{-1}]$ i $\mathbb{C}[x^{-1}, y^{-1}]$. Łatwo zobaczyć ([Ful93], przykład na s. 21), że $X_{P_{r,s}} = \mathbb{P}^1 \times \mathbb{P}^1$.



Zanurzenie tej rozmaiłości w przestrzeń rzutową zależy od kraty $M \simeq \mathbb{Z}^2$, którą dysponujemy, co wynika z rozważań podrozdziału 1.2.2. Rozważymy dla przypadku.

Przypadek pierwszy Jeżeli za M przyjmiemy kratę generowaną przez wektory $(r, 0)$ i $(0, s)$ (jej elementy to duże kropki na pierwszym rysunku), to znajdując izomorfizm krat $M \simeq \mathbb{Z}^2$ wektory $(r, 0)$ i $(0, s)$ możemy zapisać jako $(1, 0)$ i $(0, 1)$. Zbiór \mathcal{A} z równania 1.3 w będzie definiował macierz

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

a zatem w tym wypadku Y_A jest zadane jako domknięcie obrazu parametryzacji $(\mathbb{C}^*)^2 \rightarrow \mathbb{P}^3$ takiej, że $(t_1, t_2) \mapsto (1, t_1, t_2, t_1 t_2)$.

Łatwo sprawdzić, że $\ker_{\mathbb{Z}}(A) = \text{span}_{\mathbb{Z}}\{[1, -1, -1, 1]\}$, a zatem na mocy lematu 1.1 mamy $\mathcal{I}_A = \langle x_0 x_3 - x_1 x_2 \rangle$, gdzie x_i stanowią zestaw współrzędnych dla \mathbb{P}^3 .

Zauważmy, że kolumny macierzy A definiują wielościenne pierścień półgrupowy, który jest normalny, gdyż wielościan $P_{r,s}$ jest normalny względem kraty M (ma dwuelementową triangulację unimodularną).

Przypadek drugi Niech krata M będzie rozpinana przez wektory $(1, 0)$ i $(0, 1)$, to wielościan $P_{r,s}$ ma $(r+1)(s+1)$ punktów kratowych. Będziemy rozważać przypadek $r > 1$ lub $s > 1$. Jego ściany maksymalnego wymiaru zadają dywizory niezmiennicze względem działania torusa $(\mathbb{C}^*)^2$. Na przykład ściana F na pierwszym rysunku zadaje dywizor $\{t_2^s = 0\}$. Równoważna reprezentacja za pomocą półpłaszczyzn dla wielościanu $P_{r,s}$ wygląda tak:

$$P_{r,s} = \{x \geq 0\} \cap \{y \geq 0\} \cap \{-x \geq -r\} \cap \{-y \geq -s\}.$$

A zatem na mocy rozważań tego podrozdziału, $P_{r,s}$ zadaje dywizor szeroki postaci

$$D_{P_{r,s}} = rD_3 + sD_4,$$

bo $a_1 = a_2 = 0$ oraz $a_3 = r$, $a_4 = s$.

Dywizor $D_{P_{r,s}}$ wyznacza przekształcenie z $X_{P_{r,s}}$ do Y_A , gdzie zbiór \mathcal{A} z równania 1.3 w tym przypadku to

$$\{(i, j, 1) : i = 0, \dots, r \text{ oraz } j = 0, \dots, s\}.$$

Czyli $Y_A \subset \mathbb{P}^{(r+1)(s+1)-1}$. Aby łatwo zobaczyć, jak torus $T(N) = (\mathbb{C}^*)^2$ siedzi w tej rozma-
itości, można pokazać to włożenie jako odpowiednie złożenie

$$(\mathbb{C}^*)^2 \rightarrow \mathbb{P}^r \times \mathbb{P}^s \rightarrow \mathbb{P}^{(r+1)(s+1)-1}$$

$$(t_1, t_2) \mapsto (1, t_1, t_1^2, \dots, t_1^r; 1, t_2, t_2^2, \dots, t_2^s) \mapsto (t_1^i t_2^j)_{i,j} \text{ gdzie } i = 0, \dots, r; j = 0, \dots, s.$$

Na mocy zależności z poprzedniej uwagi, mamy dokładnie cztery punkty stałe działania torusa x_σ , które odpowiadają wierzchołkom wielościanu i czterem stożkom maksymalnego wymiaru $\sigma \in \Sigma_P$ – patrz [Ful93], s. 28.

Rozdział 2

Statystyka algebraiczna

W tej części pracy zaprezentujemy część teorii, w której spotyka się algebra i statystyka. Sposób prezentacji będzie możliwie zwięzły, gdyż wyniki tego podrozdziału dostarczają jedynie motywacji do rozważań zaprezentowanych w dalszej części pracy. Będziemy się koncentrować na geometrii modeli statystycznych, czyli czymś co w ramach samej statystyki algebraicznej coraz częściej nazywa się statystyką geometryczną. Poza podstawowymi intuicjami, obejmie to przede wszystkim opis takich modeli statystycznych, których analiza stała się podstawą do wykorzystania geometrii torycznej.

Na początku przedstawimy najważniejsze pojęcia statystyki oraz wybrane problemy, z którymi się ona styka. Część tych problemów, jak się później okaże, da się przetłumaczyć na język algebry i geometrii. Następnie w podrozdziale 2.2 zaprezentujemy podstawowe definicje statystyki algebraicznej oraz podamy ciekawe przykłady, jak te idee pojawiają się w prostych zagadnieniach statystycznych.

Następnie w podrozdziale 2.3 przedstawimy ciekawą rodzinę modeli statystycznych, dla których mamy w naturalny sposób zadany ideał toryczny, a zatem również rozmaitość semitoryczną. Modele te będziemy nazywać modelami torycznymi i jak się zaraz okaże, odpowiadają one statystycznym dyskretnym rodzinom wykładniczym.

Szczególnym przypadkiem modeli torycznych są modele graficzne oraz modele grafów, które zostaną przedstawione w podrozdziale 2.4. Jest to bardzo popularna rodzina modeli statystycznych, o której powstawały już całe monografie (patrz np. Whittaker [Whi90], Lauritzen [Lau96] i Edwards [Edw00]).

2.1. Model statystyczny

Mówiąc najogólniej, statystyka zajmuje się zbieraniem danych i wnioskowaniem na ich podstawie. Będziemy zajmować się przypadkiem, gdy wnioskowanie poprzedza sformułowanie pewnej wiedzy a priori w postaci modelu probabilistycznego. Zadaniem statystyki jest w tym przypadku, by uzupełnić tę wiedzę oraz zweryfikować hipotezy dotyczące jej słuszności.

Dobrym przykładem, który jednocześnie pokazuje, jaka jest intuicyjna różnica między problemami rachunku prawdopodobieństwa i statystyki jest klasyczne zadanie o urnie z kulami. W rachunku prawdopodobieństwa brzmi ono mniej więcej tak: W urnie znajduje się n kul, z czego m jest czarnych a $n - m$ białych. Jakie jest prawdopodobieństwo, że wyciągając dziesięć kul ze zwracaniem dostaniemy dokładnie 3 kule białe i 7 czarnych? W statystyce problem jest inny: Przypuśćmy, że losując 10 razy ze zwracaniem dostaliśmy 3 kule białe i 7 czarnych. Wiemy, że w urnie jest n kul. Ile w urnie jest kul białych a ile czarnych? Problem tego typu to właśnie wnioskowanie na podstawie próby.

Niech X będzie zmienną losową (lub skończonym bądź nieskończonym ciągiem zmiennych losowych) o wartościach w przestrzeni \mathcal{X} . Zmienną X będziemy często nazywać wynikiem eksperymentu lub obserwacją. Zakładamy, że \mathcal{X} jest zbiorem skończonym, przeliczalnym, lub jest podzbiorem \mathbb{R}^n i będziemy go nazywać przestrzenią obserwacji. Niech $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ będzie rodziną rozkładów prawdopodobieństwa na σ -ciele zbiorów borelowskich w \mathcal{X} (lub σ -ciele wszystkich podzbiorów \mathcal{X} , gdy jest skończony) indeksowaną przez zbiór Θ . Przestrzeń próby wraz z rodziną rozkładów \mathcal{P} , tzn.

$$\{\mathcal{X}, \{\mathbb{P}_\theta : \theta \in \Theta\}\},$$

nazywamy modelem statystycznym. Odwzorowania z \mathcal{X} do \mathbb{R}^k nazywamy statystykami. W przykładzie z urną można skonstruować następujący model statystyczny. Załóżmy, że w naszym eksperymencie będziemy dziesięć razy losować ze zwracaniem. Obserwacją X może być liczba wyciągniętych białych kul. Przestrzenią obserwacji \mathcal{X} jest zbiór $\{0, \dots, 10\}$ (przestrzeń możliwych realizacji zmiennej X). Czyli możemy jako model statystyczny dla tego doświadczenia przyjąć $\{\mathcal{X}, \{\text{Bin}(10, \theta) : \theta \in [0, 1]\}\}$. Tutaj $\{\text{Bin}(10, \theta) : \theta \in [0, 1]\}$ oznacza rodzinę rozkładów liczby sukcesów (sukces = wyciągnięcie białej kuli) w dziesięciu losowaniach, która jest parametryzowana przez prawdopodobieństwo sukcesu (wyciągnięcia białej kuli) w pojedynczym losowaniu. Jest to tak zwana rodzina rozkładów dwumianowych o parametrach 10 i $\theta \in [0, 1]$ (patrz przykład 2.1).

Podstawowe zadanie związane z analizą statystyczną wygląda mniej więcej tak: Przypuśćmy, że dany jest model statystyczny $\{\mathcal{X}, \{\mathbb{P}_\theta : \theta \in \Theta\}\}$ oraz k obserwacji X_1, \dots, X_k o wartościach w \mathcal{X} , które są niezależne i pochodzą z rozkładu \mathbb{P}_θ (mówimy wtedy o próbie prostej). Najogólniej mówiąc, zadanie polega by na podstawie tej próby dowiedzieć się jak najwięcej o nieznanym parametrze θ . Mniej formalnie: polega to na powtórzeniu k niezależnych eksperymentów dla badanej zmiennej oraz wnioskowaniu na tej podstawie na jej temat. W przykładzie z kulami pojedynczy eksperyment polega na wykonaniu dziesięciu niezależnych losowań ze zwracaniem i zanotowaniu liczby wyciągniętych kul białych. Eksperyment taki pozwala wnioskować o nieznanym prawdopodobieństwie wyciągnięcia kuli białej, jednak nasze wyniki uwiarygodniają się, gdy liczba eksperymentów jest większa.

Pytanie o konkretną wartość θ związane jest z zagadnieniem estymacji. Formalnie polega to na skonstruowaniu przekształcenia $\hat{\theta} : \mathcal{X}^k \rightarrow \Theta$, które będzie w odpowiedni sposób przybliżać prawdziwą wartość θ . Wracając do naszego przykładu: Jeżeli przez x_i dla $i = 1, \dots, k$ będziemy oznaczać zanotowaną liczbę kul białych dla i -tego eksperymentu spośród k przeprowadzonych eksperymentów (każde po dziesięć losowań), to powszechnie stosowanym estymatorem parametru θ jest $\hat{\theta} = \frac{1}{k} \sum_{i=1}^k \frac{x_i}{10}$. W szczególności przy pojedynczym eksperymencie „najlepszym”¹ estymatorem parametru θ jest częstość wyciągania białych kul $\frac{x}{10}$.

Innym typem problemu jest weryfikacja twierdzenia $\theta \in \Theta_0$ kontra $\theta \in \Theta_1$, gdzie Θ_0 i Θ_1 są dwoma rozłącznymi podzbiórmi zbioru Θ . Zadania tego typu nazywamy weryfikacją hipotez. Typowy przykład to weryfikacja hipotezy $p = \frac{1}{2}$ przeciw $p \neq \frac{1}{2}$, gdzie p to prawdopodobieństwo wyrzucenia orła, na podstawie wyników próby złożonej ze 100 rzutów monetą. W tym wypadku weryfikujemy hipotezę mówiącą, że moneta, którą rzucamy jest symetryczna. Należy oczekiwać, że jeśli liczba wyrzuconych orłów znacznie różni się od pięćdziesięciu, to przeczy to naszej hipotezie o symetryczności monety. Weryfikacja hipotez jest związana również z ważną charakterystyką statystyki jako nauki zajmującej się modelami, podmodelami i relacjami między nimi (patrz [PRW01], s. 120), która jest szczególnie ważna w kontekście wykorzystania geometrii algebraicznej. Podmodel modelu $\{\mathcal{X}, \mathcal{P}\}$ to przestrzeń obserwacji wraz z pewną podrodziną rodziny \mathcal{P} . Będziemy jeszcze pisać o tym później.

¹Pożądane własności, które powinien spełniać dobry estymator opisane są na przykład w książce Zielińskiego [Zie90].

Wymienione zagadnienia są jedynie częścią nowoczesnej statystyki, którą zwie się statystyką klasyczną. Nam jednak one wystarczą, jako ilustracje związków pomiędzy algebrą i geometrią algebraiczną a statystyką. Wiele ważnych zagadnień nawet w ramach samej estymacji i weryfikacji hipotez będziemy jednak musieli pominąć, czego niestety wymaga ograniczona forma tej pracy. Pominiemy również wiele ciekawych zagadnień rozwijanych w ramach statystyki algebraicznej koncentrując się prawie wyłącznie na tych najmocniej związanych z wykorzystaniem geometrii torycznej. Statystyka algebraiczna w szerszym kontekście przedstawiona jest w monografii [PRW01] oraz książce [SP05].

2.2. Geometria modeli statystycznych

W pracy tej będziemy zakładać, że przestrzeń obserwacji jest skończona² $|\mathcal{X}| = m$. W takim przypadku możemy identyfikować przestrzeń stanów \mathcal{X} z pierwszymi m liczbami naturalnymi

$$[m] := \{1, 2, \dots, m\}.$$

Niech X będzie obserwacją o wartościach w \mathcal{X} . Rozkład prawdopodobieństwa zmiennej losowej X , czyli m -elementowy ciąg

$$(\mathbb{P}(X = i))_{i=1}^m \text{ taki, że } \sum_{i=1}^m \mathbb{P}(X = i) = 1,$$

możemy utożsamić z punktem na $(m - 1)$ -wymiarowym simpleksie probabilistycznym

$$\Delta_{m-1} := \{(p_1, p_2, \dots, p_m) \in \mathbb{R}^m : \sum_{i=1}^m p_i = 1 \text{ oraz } p_j \geq 0 \text{ dla każdego } j\}.$$

Wówczas i -ta współrzędna p_i odpowiada prawdopodobieństwu zaobserwowania stanu i . Zatem możemy równoważnie powiedzieć, że model statystyczny \mathcal{M}_{m-1} to pewien podzbiór simpleksu Δ_{m-1} (podzbiór ten utożsamiamy z rodziną rozkładów). W przypadku skończonej przestrzeni stanów w sposób bezpośredni dostajemy zatem przedstawienie geometryczne modelu statystycznego.

Przykład 2.1 (Rozkład dwumianowy) *Przypuśćmy, że przeprowadzamy m razy ten sam eksperyment, którego wyniki Y_1, \dots, Y_m mają charakter binarny, tzn. $Y_i \in \{0, 1\}$. Jeśli $Y_i = 1$, to mówimy, że w i -tej próbie zanotowaliśmy sukces. Możemy na przykład m razy rzucać monetą i o sukcesie mówić za każdym razem, gdy wyrzucono reszkę.*

Niech zmienna losowa X będzie liczbą sukcesów w m eksperymentach. Wówczas $\mathcal{X} = \{0, 1, \dots, m\}$. Jeśli prawdopodobieństwo sukcesu w eksperymencie wynosi θ (tzn. $\mathbb{P}(Y_1 = 1) = \dots = \mathbb{P}(Y_m = 1) = \theta$), to rozkład prawdopodobieństwa na \mathcal{X} ma postać:

$$\mathbb{P}(X = j) = \binom{m}{j} \theta^j (1 - \theta)^{m-j}, \quad \theta \in [0, 1], \quad j = 0, \dots, m.$$

Zatem model statystyczny jako rodzina odpowiednich rozkładów prawdopodobieństwa ma postać:

$$\mathcal{M}_m = \left\{ \left((1 - \theta)^m, \binom{m}{1} \theta (1 - \theta)^{m-1}, \binom{m}{2} \theta^2 (1 - \theta)^{m-2}, \dots, \theta^m \right) \mid \theta \in [0, 1] \right\},$$

czyli geometrycznie jest krzywą gładką w standardowym simpleksie m -wymiarowym parametryzowaną przez odcinek $[0, 1]$.

²Statystyka algebraiczna jest coraz szerszym pojęciem i istnieje wiele jej zagadnień, w których takiego założenia się nie czyni.

Istotnym faktem motywującym dalsze działania jest, że podzbiór $\mathcal{M} \subset \Delta$ często w praktyce daje się opisać jako przecięcie Δ ze zbiorem zer wielomianów. Jak zaraz się okaże, tak w istocie dzieje się z modelem z powyższego przykładu. Motywuje to następującą definicję.

Definicja 2.2 Niech $S \subseteq \mathbb{R}[\mathbf{p}] := \mathbb{R}[p_1, \dots, p_m]$ będzie dowolnym podzbiorem pierścienia wielomianów. Oraz niech $V_{\mathbb{R}}(S) = \{\mathbf{a} \in \mathbb{R}^m \mid f(\mathbf{a}) = 0 \quad \forall f \in S\}$ będzie odpowiadającą mu rozmaitością algebraiczną. Wówczas przecięcie:

$$V_{\Delta}(S) = V_{\mathbb{R}}(S) \cap \Delta_{m-1}$$

czyli rodzinę rozkładów prawdopodobieństwa zadaną przez warunki algebraiczne będziemy nazywać algebraicznym modelem statystycznym.

Przykład 2.3 (Rozkład dwumianowy c.d.) Jeżeli \mathcal{M}_m będzie jak w przykładzie 2.1 to zbiór S_m będzie zdefiniowany jako zbiór wszystkich 2×2 -minorów macierzy:

$$\begin{bmatrix} p_0 & p_1/\binom{m}{1} & p_2/\binom{m}{2} & \cdots & p_{m-1}/\binom{m}{m-1} \\ p_1/\binom{m}{1} & p_2/\binom{m}{2} & p_3/\binom{m}{3} & \cdots & p_m \end{bmatrix}$$

I zachodzi równość $\mathcal{M}_m = V_{\Delta}(S_m)$.

W poprzednim podrozdziale nic nie powiedzieliśmy o zbiorze parametrów Θ . W praktyce z reguły liczą się tylko tak zwane parametryczne modele statystyczne, które parametryzowane są przez skończenie wymiarowy zbiór Θ .

Definicja 2.4 Niech $\Theta \subset \mathbb{R}^d$ będzie przestrzenią parametrów a ϕ przekształceniem ciągłym $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^m$ takim, że $\phi(\Theta) \subseteq \Delta_{m-1}$. Wówczas $\phi(\Theta)$ nazywamy parametrycznym modelem statystycznym.

W praktyce będziemy zakładać nie tylko $d < \infty$ lecz również $d \ll m$.

Aby przejść do algebraicznych parametrycznych modeli statystycznych, musimy założyć, że Θ jest zbiorem semialgebraicznym, tzn. jest zbiorem rozwiązań skończonej liczby równań i nierówności wielomianowych (patrz Sullivant [Sul05], s. 5).

Definicja 2.5 Niech $\Theta \subset \mathbb{R}^d$ będzie zbiorem semialgebraicznym oraz ϕ przekształceniem wielomianowym

$$\phi : \mathbb{R}^d \rightarrow \mathbb{R}^m, \quad \theta = (\theta_1, \dots, \theta_d) \mapsto (\phi_1(\theta), \dots, \phi_m(\theta))$$

takim, że $\phi_i = \frac{f_i}{g_i}$ dla $f_i, g_i \in \mathbb{R}[\theta_1, \dots, \theta_d]$ oraz zachodzi $\phi(\Theta) \subseteq \Delta_{m-1}$. Wówczas domknięcie obrazu $\phi(\Theta)$ nazywamy algebraicznym parametrycznym modelem statystycznym.

Uwaga 2.6 W rozważaniach geometrycznych będziemy rozważać parametryzację nad ciałem liczb zespolonych. Będziemy w tej sytuacji używać tego samego oznaczenia ϕ .

Ograniczenie się do algebraicznych parametrycznych modeli statystycznych pozwala zachować bardzo dużą ogólność rozważań. Zdecydowana większość dyskretnych modeli statystycznych stosowanych w praktyce daje się przedstawić w tej postaci. Jeżeli dla jakiejś klasy modeli da się znaleźć przedstawienie ich parametryzacji w postaci wielomianowej lub wymiernej, stwarza to możliwość zarówno wykorzystania narzędzi algebry przemiennej, jak i badania geometrii zbiorów, które powstają w tym kontekście.

Ciężko w tej chwili jednoznacznie stwierdzić na ile geometria algebraiczna, badając geometrię modeli parametrycznych, może wspomóc analizę statystyczną. Z pewnością jest wiele

prostych ale interesujących pytań, na które geometria algebraiczna potrafi odpowiedzieć. Na przykład wymiar rozmaitości związanej z modelem statystycznym, to minimalna liczba parametrów potrzebnych do jego opisu. Innym przykładem może być szczegółowa analiza podmodeli modelu statystycznego, co w języku geometrii odpowiada badaniu podrozmaitości rozmaitości algebraicznej. W kontekście modeli niezależności omówimy ten przykład w następnym podrozdziale. Czasem mamy też do czynienia z sytuacją, gdy zadana mamy rozmaitość związaną z modelem i zastanawiamy się nad odwrotną parametryzacją, czyli nad znalezieniem odpowiedniego algebraicznego modelu statystycznego. W tym kontekście statystyka geometryczna może w przyszłości pomóc odpowiedzieć na wiele ważnych pytań dotyczących identyfikacji modeli statystycznych, czyli na przykład, kiedy dwa różne modele zadają taką samą rozmaitość algebraiczną. Więcej o zastosowaniach geometrii algebraicznej w kontekście algebraicznych modeli statystycznych można znaleźć na przykład w świeżym opracowaniu Drtona i Sulivanta [DS06].

Wszystkie te zagadnienia są stosunkowo proste. Statystyka geometryczna w swoich rozważaniach jest już jednak znacznie dalej (patrz na przykład [BW06]) i zajmuje się nierzadko bardzo abstrakcyjnymi własnościami odpowiednich rozmaitości. Wydaje się jednak, że nawet z punktu widzenia statystyki ważne, by badania te były kontynuowane. W tej chwili bowiem przede wszystkim zbierane jest jak najwięcej geometrycznej informacji o tych rozmaitościach, lecz w krótkiej przyszłości z pewnością prowadzone będą prace, mające jeszcze mocniej osadzić te rozważania w analizie statystycznej. Przykładem, kiedy rozważania geometryczne w większy sposób łączą się z analizą statystyczną, są modele drzew filogenetycznych w biologii (patrz na przykład [ERSS04]).

2.3. Modele toryczne i modele niezależności

W statystyce ważną klasę modeli stanowią rodziny wykładnicze (patrz [Zie90], s. 28). W przypadku dyskretnym rozważa się rozkłady prawdopodobieństwa postaci:

$$P_\theta(x) = Z(\theta) \cdot e^{\langle \theta, T(x) \rangle}, \quad \theta \in [-\infty, \infty)^d, \quad (2.1)$$

gdzie $x \in \mathcal{X} = [m]$, $Z(\theta)$ jest stałą normującą, $\langle \cdot, \cdot \rangle$ oznacza iloczyn wewnętrzny a $T : \mathcal{X} \rightarrow \mathbb{Z}^d \setminus \{0\}$ jest d -wymiarową statystyką (to znaczy odwzorowaniem \mathcal{X} w jakiś podzbiór \mathbb{R}^d).

Na pierwszy rzut oka nie widać, jak dyskretne rodziny wykładnicze wiążą się z modelami algebraicznymi. Jak się zaraz okaże przez prostą zamianę zmiennych możemy uzyskać parametryzację algebraiczną dla tych modeli. Wygodne będzie następujące, bardzo popularne (patrz Geiger i in. [GMS02]) przeformułowanie tych modeli. Niech $A = (a_{ij})$ będzie macierzą wymiaru $d \times m$ o elementach będących liczbami naturalnymi, taką, że sumy elementów w każdej kolumnie są równe³. Wówczas j -ta kolumna a_j reprezentuje jednomian $t^{a_j} = \prod_{i=1}^d t_i^{a_{ij}}$ oraz wszystkie jednomiany są tego samego stopnia. Z macierzą A możemy związać zatem przekształcenie wielomianowe:

$$\phi_A : \mathbb{R}^d \rightarrow \mathbb{R}^m, \quad (t_1, \dots, t_d) \mapsto \left(\prod_{i=1}^d t_i^{a_{i1}}, \dots, \prod_{i=1}^d t_i^{a_{im}} \right). \quad (2.2)$$

Model statystyczny związany z powyższą parametryzacją to para $\{\mathcal{X}, \mathcal{P}\}$, gdzie $\mathcal{X} = [m]$ oraz $\mathcal{P} = \{\mathbb{P}_t | t \in \mathbb{R}^d\}$, gdzie $\mathbb{P}_t = \phi_A(t) \cap \Delta_{m-1}$. Będziemy go nazywać modelem torycznym.

³Jak się zaraz okaże, zbieżność z notacją z poprzedniego rozdziału nie jest przypadkowa. Dodatkowo założenie, że sumy elementów w każdej kolumnie są równe oznacza w szczególności, że wektor jedynek leży w przestrzeni rozpinanej przez wiersze macierzy A .

Nazwa bierze się stąd, że domknięcie tego obrazu definiuje obiekt analogiczny do rozmaitości semitorycznej, tylko zdefiniowany nad ciałem liczb rzeczywistych. Konstrukcja ta została przedstawiona szczegółowo w podrozdziale 1.1.

Zauważmy, że klasa modeli statystycznych z równania 2.1 oraz równania 2.2 jest identyczna. Można zapomnieć o stałej normującej pamiętając o tym, że obraz $\mathbb{P}_\theta(\cdot)$ leży w simpleksie probabilistycznym. Każda kolumna A odpowiada innemu stanowi przestrzeni \mathcal{X} . Model zdefiniowany przez 2.1 ze statystyką $T(x)$ jest równoważny modelowi 2.2 z macierzą A wtedy i tylko wtedy, gdy kolumny a_j macierzy A równe są odpowiednim $T(x)$. Przejście między modelami polega na prostej zamianie parametrów θ_i na t_i poprzez przekształcenie $t_i = e^{\theta_i}$, gdzie $e^{-\infty} = 0$.

Nie będziemy się zajmować algebraicznymi modelami statystycznymi aż w takiej ogólności, koncentrując się jedynie pewnymi podrodzinami dyskretnych rodzin wykładniczych. Jednak wyjście od tak szerokiej klasy modeli, jak modele toryczne, pokazuje stosunkowo dużą ogólność rozważań, którą możemy zachować ograniczając się w analizie statystycznej skończonych przestrzeni obserwacji do algebraicznych modeli statystycznych.

Często w praktyce badając system złożony z wielu obserwowanych zmiennych możemy zastanawiać się nad zależnościami i niezależnościami między poszczególnymi zmiennymi. Przyjmijmy na początek, że obserwujemy dwie dyskretne zmienne losowe X_1 oraz X_2 o wartościach odpowiednio w zbiorach $[m_1]$ oraz $[m_2]$. Ich łączny rozkład prawdopodobieństwa możemy utożsamić z macierzą $P = [p_{ij}] = \mathbb{P}(X_1 = i, X_2 = j)$ wymiaru $m_1 \times m_2$. Zauważmy, że $p_{ij} \geq 0$ oraz $\sum_{i,j} p_{ij} = 1$, a zatem $P \in \Delta_{m_1 m_2 - 1}$. Dodatkowo $(p_{i+})_{i \in [m_1]}$, gdzie $p_{i+} = \sum_{j=1}^{m_2} p_{ij}$, leży w simpleksie $(m_1 - 1)$ -wymiarowym $\Delta_{m_1 - 1}$. Analogicznie $(p_{j+})_{j \in [m_2]}$ leży w simpleksie $(m_2 - 1)$ -wymiarowym $\Delta_{m_2 - 1}$.

Definicja 2.7 *Dwie dyskretne zmienne losowe X_1 i X_2 są niezależne wtedy i tylko wtedy, gdy $p_{ij} = \mathbb{P}(X_1 = i, X_2 = j) = \mathbb{P}(X_1 = i) \cdot \mathbb{P}(X_2 = j) = p_{i+} \cdot p_{+j}$ dla wszystkich $(i, j) \in [m_1] \times [m_2]$.*

W przypadku skończonych przestrzeni stanów badanie niezależności dwóch zmiennych sprowadza się do sprawdzenia prostego algebraicznego warunku na ich rozkład łączny.

Lemat 2.8 *Dwie zmienne losowe X_1 i X_2 są niezależne wtedy i tylko wtedy, gdy macierz ich rozkładu łącznego $P = [p_{ij}]$ ma rząd 1.*

Dowód. Macierz P wymiaru $m_1 \times m_2$ ma rząd 1 wtedy i tylko wtedy, gdy istnieją dwa wektory (u_1, \dots, u_{m_1}) oraz (v_1, \dots, v_{m_2}) o współczynnikach rzeczywistych takie, że $p_{ij} = u_i v_j$ dla wszystkich i oraz j . Możemy unormować elementy wektora u tak, aby spełnione było $\sum_i u_i = 1$. Dostajemy:

$$v_j = \left(\sum_i u_i \right) \cdot v_j = \sum_i p_{ij} = p_{+j}.$$

Zakładając, że P leży w $\Delta_{m_1 m_2 - 1}$, otrzymujemy

$$u_i = u_i \cdot \left(\sum_j p_{+j} \right) = \sum_j u_i v_j = p_{i+}.$$

Zatem macierz rozkładu łącznego P ma rząd 1 wtedy i tylko wtedy, gdy

$$p_{ij} = p_{i+} \cdot p_{+j} \text{ dla wszystkich } (i, j) \in [m_1] \times [m_2],$$

co jest dokładnie definicją niezależności zmiennych X_1 i X_2 . \square

Statystyczny model niezależności $\mathcal{M}_{m_1 m_2 - 1}$ jest modelem torycznym dla $d = m_1 + m_2$ oraz $m = m_1 m_2$. Punkt $p \in \Delta_{m_1 m_2 - 1}$ leży w obrazie przekształcenia

$$f : \mathbb{R}^{m_1 + m_2} \rightarrow \mathbb{R}^{m_1 m_2}, \quad (\theta_1, \dots, \theta_{m_1 + m_2}) \mapsto (\theta_i \theta_{j+m_1})_{i \in [m_1], j \in [m_2]}$$

wtedy i tylko wtedy, gdy X_1 oraz X_2 są niezależne; czyli wtedy i tylko wtedy, gdy macierz rozkładu łącznego ma rząd 1.

Z powyższego rozumowania łatwo wynika postać macierzy A , która definiuje model toryczny. Na przykład dla $m_1 = 2$ oraz $m_2 = 3$ mamy.

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \quad (2.3)$$

I macierz A indukuje przekształcenie

$$f : \mathbb{R}^5 \rightarrow \mathbb{R}^{2 \times 3}, \quad (\theta_1, \dots, \theta_5) \mapsto \begin{bmatrix} \theta_1 \theta_3 & \theta_1 \theta_4 & \theta_1 \theta_5 \\ \theta_2 \theta_3 & \theta_2 \theta_4 & \theta_2 \theta_5 \end{bmatrix}.$$

Można pokazać⁴, że ideał wszystkich wielomianów zerujących się na statystycznym modelu niezależności \mathcal{M}_m generowany jest przez podwyznaczniki wymiaru 2×2 macierzy P , tzn

$$\mathcal{I}(\mathcal{M}_m) = \langle p_{ij} p_{kl} - p_{il} p_{kj} \mid 1 \leq i, k \leq m, 1 \leq j, l \leq d \rangle.$$

Uwaga 2.9 *Zauważmy, że w powyższym przykładzie pominęliśmy założenie, że wielomian $\sum_{i,j} p_{ij} - 1$ leży w ideale $\mathcal{I}(\mathcal{M}_m)$. Warunek ten jest nazywany trywialnym niezmiennikiem i jest często pomijany w statystyce algebraicznej. Zabieg ten jest celowy i pozwala rozważać $\mathcal{M}_m \cap \Delta_{m_1 m_2}$ jako nieujemną część obrazu Segre przestrzeni $\mathbb{P}_{\mathbb{R}}^{m_1 - 1} \times \mathbb{P}_{\mathbb{R}}^{m_2 - 1}$ w przestrzeni $\mathbb{P}_{\mathbb{R}}^{m_1 m_2 - 1}$. W ogólności zaś pozwala myśleć o modelach statystycznych jako rozmaitościach leżących w przestrzeni rzutowej (patrz [Sul05], s. 6). Zgodnie z uwagą 2.6, w rozważaniach geometrycznych za ciało bazowe będziemy dla uproszczenia brać ciało liczb zespolonych.*

Dla większej liczby obserwowanych zmiennych mamy dużo więcej możliwych modeli odwołujących się do niezależności i niezależności warunkowych. Rzeczywiście, już dla trzech zmiennych możemy rozważać na przykład niezależność dwóch z nich pod warunkiem trzeciej. Niech X , Y i Z będą trzema dyskretnymi zmiennymi losowymi o wartościach odpowiednio w \mathcal{X} , \mathcal{Y} i \mathcal{Z} . Mówimy, że X jest warunkowo niezależny od Y pod warunkiem Z i piszemy $X \perp\!\!\!\perp Y \mid Z$, jeśli:

$$\mathbb{P}(X = x, Y = y \mid Z = z) = \mathbb{P}(X = x \mid Z = z) \cdot \mathbb{P}(Y = y \mid Z = z),$$

gdzie równanie zachodzi dla wszystkich $x \in \mathcal{X}$, $y \in \mathcal{Y}$ oraz dla wszystkich $z \in \mathcal{Z}$ spełniających $\mathbb{P}(Z = z) > 0$. Równoważnie

$$\mathbb{P}(X = x, Y = y, Z = z) = \frac{\mathbb{P}(X = x, Z = z) \cdot \mathbb{P}(Y = y, Z = z)}{\mathbb{P}(Z = z)}, \quad (2.4)$$

⁴Na przykład stosując technologię baz Gröbnera, jak to zostało opisane w artykule Diaconisa i Sturmfelsa [DS98]

czyli prawdopodobieństwo łączne zmiennych X, Y, Z faktoryzuje się odpowiednio co parametrycznie możemy zapisywać jako

$$p_{xyz} = t_{xz}^{XZ} \cdot t_{yz}^{YZ}, \quad (2.5)$$

co ma być spełnione dla wszystkich możliwych $x \in \mathcal{X}$, $y \in \mathcal{Y}$ i $z \in \mathcal{Z}$, co daje $|\mathcal{X}| \cdot |\mathcal{Y}| \cdot |\mathcal{Z}|$ parametrów typu p_{xyz} , $|\mathcal{X}| \cdot |\mathcal{Z}|$ parametrów typu t_{xz}^{XZ} oraz $|\mathcal{Y}| \cdot |\mathcal{Z}|$ parametrów typu t_{yz}^{YZ} .

To co jest najważniejsze, to fakt, że parametryzacja 2.5 zadaje przekształcenie wielomianowe analogiczne do tego w równaniu 2.2. Jeżeli będziemy rozważać całą sytuację nad ciałem liczb zespolonych, to w konsekwencji zadaje ona również rozmaitość semitoryczną. W przypadku trzech zmiennych mamy do czynienia z większą liczbą modeli zapisywanych języku warunkowych niezależności. Każdy z takich modeli wyznacza inną parametryzację, a zatem odpowiada innej rozmaitości w simpleksie probabilistycznym. Często w statystyce rozważając model, bada się również jego podmodele. Na przykład model $\mathcal{M}_0 : X \perp\!\!\!\perp Y | Z$, którego rozkłady łączne spełniają warunek z równania 2.4, zawiera model $\mathcal{M}_1 : X \perp\!\!\!\perp (Y, Z)$. Rzeczywiście rozkłady modelu \mathcal{M}_1 spełniają:

$$\mathbb{P}(X = x, Y = y, Z = z) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y, Z = z) \quad (2.6)$$

dla wszystkich x, y, z (z definicji niezależności zmiennej X od pary zmiennych (Y, Z)). W szczególności zatem

$$\sum_{y \in \mathcal{Y}} \mathbb{P}(X = x, Y = y, Z = z) = \sum_{y \in \mathcal{Y}} \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y, Z = z).$$

Ale zachodzą oczywiste równości

$$\sum_{y \in \mathcal{Y}} \mathbb{P}(X = x, Y = y, Z = z) = \mathbb{P}(X = x, Z = z)$$

$$\sum_{y \in \mathcal{Y}} \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y, Z = z) = \mathbb{P}(X = x) \cdot \mathbb{P}(Z = z),$$

czyli $\mathbb{P}(X = x, Z = z) = \mathbb{P}(X = x) \cdot \mathbb{P}(Z = z)$. Możemy wobec tego w równaniu 2.6 zamiast $\mathbb{P}(X = x)$ wstawić $\mathbb{P}(X = x, Z = z) / \mathbb{P}(Z = z)$. Czyli

$$\mathbb{P}(X = x, Y = y, Z = z) = \frac{\mathbb{P}(X = x, Z = z) \cdot \mathbb{P}(Y = y, Z = z)}{\mathbb{P}(Z = z)},$$

co jest dokładnie równaniem 2.4. Czyli wszystkie rozkłady spełniające $X \perp\!\!\!\perp (Y, Z)$ spełniają również $X \perp\!\!\!\perp Y | Z$. Zatem parametryzacja modelu \mathcal{M}_1 wyznacza podrozmaitość \mathcal{V}_1 rozmaitości \mathcal{V}_0 parametryzowanej przez \mathcal{M}_0 . Jest to ciekawa geometryczna interpretacja podmodelu modelu statystycznego. I tak na przykład rozważanie zbioru otwartego $\mathcal{V}_0 \setminus \mathcal{V}_1$ pozwala odpowiedzieć na pytanie „jak dużo” jest rozkładów prawdopodobieństwa, które spełniają formułę \mathcal{M}_0 a nie spełniają \mathcal{M}_1 . Może prowadzić to zarówno do ciekawych problemów praktycznych, jak i teoretycznych.

2.4. Modele graficzne i modele grafów

Modele niezależności z reguły formułuje się dla systemów, w których obserwuje się więcej zmiennych losowych. W tym przypadku analiza staje się stosunkowo skomplikowana i generowanie odpowiednich parametryzacji statystycznych modeli algebraicznych, jak w równaniu 2.5 staje się trudniejsze. Ciekawy wynik z pracy Geigera i Pearla [GP93] pokazuje, że

w pewnym sensie całkowitą informację o formułach niezależności warunkowej spełnionych przez układ zmiennych losowych można zapisać w grafie nieskierowanym. Jest to powód, dla którego statystyczne modele graficzne stały się w ostatnich latach tak popularnym narzędziem modelowania warunkowych niezależności. Załóżmy, że n -elementowy zbiór wierzchołków V pewnego grafu $\mathcal{G} = (V, E)$ odpowiada zmiennym losowym w analizowanym systemie $X_V = (X_1, \dots, X_n)$ o wartościach w $\mathcal{X} = \times_{i=1}^n \mathcal{X}_i$. Niech zbiór krawędzi E grafu \mathcal{G} odpowiada pewnym niezależnościom warunkowym pomiędzy zmiennymi lub ich grupami (czyli rozłącznymi podzbiórmi X_V). Wówczas - mało formalnie - modelem graficznym będziemy nazywać trójkę $\{\mathcal{X}, \mathcal{G}, \mathcal{P}_{\mathcal{G}}\}$, gdzie \mathcal{G} jest grafem o zbiorze wierzchołków V , a $\mathcal{P}_{\mathcal{G}}$ jest rodziną rozkładów prawdopodobieństwa dla X_V , które spełniają formuły niezależności warunkowych, które reprezentowane są przez zbiór krawędzi E .

W stosowanej w tym podrozdziale notacji, będziemy się wzorować w dużym stopniu na pracy Hoştena i Sullivanta [HS02]. Ogólnie rozważamy zbiór dyskretnych zmiennych losowych $X_V = \{X_1, \dots, X_n\}$, gdzie X_i przyjmuje wartości w $[d_i]$, niech $d = (d_1, \dots, d_n)$. Rozpatrujemy tablice n -wymiarowe, których elementy są indeksowane zbiorem $D = [d_1] \times \dots \times [d_n]$. Są one zatem elementami przestrzeni $\mathbb{R}^{\prod_i d_i} = \mathbb{R}^{|D|} = \otimes_{i=1}^n \mathbb{R}^{d_i}$. Współrzędne przestrzeni $\mathbb{R}^{|D|}$ będziemy oznaczać przez x_I , a jej wektory jednostkowe przez e_I , gdzie $I = (i_1, \dots, i_n) \in [d_1] \times \dots \times [d_n]$. Będziemy stosować również często zapis tensorowy, który okaże się wygodny w rozważaniach geometrycznych w trzecim rozdziale tej pracy.

W dalszej części tego rozdziału ograniczymy się do przypadku binarnego, który jest przedmiotem naszego zainteresowania w rozdziale trzecim. W przypadku binarnym X_i przyjmują wartości w zbiorze $\{0, 1\}$ i $D = \{0, 1\}^n$, a zatem $\mathbb{R}^{|D|} = \mathbb{R}^{2^n} = \otimes_{i=1}^n \mathbb{R}^2$. Wektory bazowe przestrzeni $\otimes_{k=1}^n \mathbb{R}^2$ indeksowane są w naturalny sposób ciągami zero-jedynkowymi. Będziemy je porządkować leksykograficznie zakładając $0 >_{\text{lex}} 1$. Jeżeli na przykład założymy, że $n = 4$, to mamy $0000 >_{\text{lex}} 0001 >_{\text{lex}} 0010 >_{\text{lex}} \dots >_{\text{lex}} 1111$.

Będziemy rozważać zbiór zmiennych indeksowanych przez dowolny zbiór F

$$X_F = \{X_k | k \in F, F \subset [n]\},$$

a przez i_F będziemy oznaczać realizację X_F , tzn. $i_F = (X_F(\omega)) \in \{0, 1\}^{|F|}$ dla $\omega \in \Omega$, gdzie Ω jest przestrzenią zdarzeń elementarnych. Wektory bazowe przestrzeni $\mathbb{R}^{2^{|F|}} = \otimes_{i=1}^{|F|} \mathbb{R}^2$ również indeksowane są ciągami zero-jedynkowymi i także w tym przypadku będziemy stosować porządek leksykograficzny. W naszej pracy ograniczyliśmy się do przypadku binarnego, zatem przez model graficzny będziemy rozumieć⁵: graf nieskierowany \mathcal{G} o wierzchołkach indeksowanych zbiorem $[n]$ wraz z rodziną rozkładów na przestrzeni stanów $\mathcal{X} = D = \{0, 1\}^n$ spełniających

$$\mathbb{P}(X_1 = i_1, X_2 = i_2, \dots, X_n = i_n) = \frac{1}{Z} \prod_{F \in \mathcal{C}} \Psi_F(i_F), \quad (2.7)$$

dla pewnych funkcji nieujemnych $\Psi_F : \mathcal{X} \rightarrow \mathbb{R}$, gdzie Z jest stałą normującą, a \mathcal{C} jest zbiorem klik grafu \mathcal{G} , czyli maksymalnych podgrafów pełnych grafu \mathcal{G} (tzn. między każdą parą wierzchołków w tym podgrafie jest krawędź). Model graficzny będziemy zapisywać poprzez podanie zbioru \mathcal{C} .

Pokażemy teraz, jak powyższą definicję połączyć z nieformalną definicją zaproponowaną na początku tego podrozdziału. W szczególności pokażemy w jaki sposób zbiór krawędzi grafu ma reprezentować formuły niezależności spełniane przez wszystkie rozkłady w modelu. Niech graf $\mathcal{G} = \{E, V = [n]\}$ będzie zdefiniowany tak, że brak krawędzi między wierzchołkiem α a

⁵Modele graficzne są szczególnym przypadkiem modeli hierarchicznych, które w kontekście algebraicznym zostały przedstawione na przykład przez Sullivanta [Sul05].

β w grafie \mathcal{G} oznacza, że zmienne losowe X_α i X_β są niezależne pod warunkiem wszystkich innych zmiennych w systemie, tzn.

$$(\alpha, \beta) \notin E \iff X_\alpha \perp\!\!\!\perp X_\beta | \text{reszta.}$$

Z drugiej strony, jeśli w modelu spełnione jest formuła

$$X_\alpha \perp\!\!\!\perp X_\beta | \text{reszta}$$

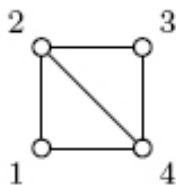
to prawdopodobieństwo łączne z definicji niezależności warunkowej faktoryzuje się tak, że w żadnym z czynników nie istnieje zarówno razem X_α i X_β . Istotnie, wystarczy założyć, że $\alpha = 1$ i $\beta = 2$, wówczas:

$$\mathbb{P}(X_V) = \mathbb{P}(X_1, X_2 | Z) \cdot \mathbb{P}(Z) = \mathbb{P}(X_1 | Z) \cdot \mathbb{P}(X_2 | Z) \cdot \mathbb{P}(Z) = \frac{\mathbb{P}(X_1, Z) \cdot \mathbb{P}(X_2, Z)}{\mathbb{P}(Z)},$$

gdzie $Z = (X_3, \dots, X_n)$. Stosując tę procedurę iteracyjnie jesteśmy w stanie rozbić prawdopodobieństwo łączne na czynniki, które będą zależeć jedynie od wartości takich grup zmiennych losowych, że każda para zmiennych losowych w grupie jest zależna pod warunkiem wszystkich innych zmiennych w systemie. Na mocy definicji grafu \mathcal{G} są to dokładnie grupy zmiennych, które indukują maksymalne podgrafy pełne w tym grafie. Pokazaliśmy, że istotnie zbiór krawędzi tego grafu zadaje warunki niezależności, które mają być spełnione i jednocześnie rodzinę rozkładów probabilistycznych, która je spełnia. Z tym podstawowych relacji niezależności typu $X \perp\!\!\!\perp Y | \text{reszta}$ można również odczytać wiele innych. Mówią o tym wzajemna, lokalna i globalna własność Markowa (patrz [Lau96], s. 32). Zagadnienia te wychodzą jednak zdecydowanie poza zakres tej pracy i nie będą szerzej omawiane.

Przykład 2.10 Niech X_1, X_2, X_3, X_4 będą czterema binarnymi obserwacjami. Analizujemy odpowiedzi na dwa pytania w ankiecie w zależności od pewnych charakterystyk respondenta. Niech X_1 , oznacza odpowiedź na pytanie „Czy oglądasz serial 'M jak miłość'?", gdzie odpowiedź „nie” traktujemy zero, a „tak” jako jeden. X_3 to odpowiedź na pytanie „Czy przeczytałaś/eś 'Ulissesa' (w całości)?”. X_2 to płeć respondenta, a X_4 to jego wykształcenie, gdzie za zero przyjmujemy: brak lub podstawowe, a za jeden: średnie i wyższe. Załóżmy, że z praktycznego punktu widzenia najbardziej interesująca była analiza modelu $X_1 \perp\!\!\!\perp X_3 | X_2, X_4$ czyli inaczej $X_1 \perp\!\!\!\perp X_3 | \text{reszta}$. Naturalne jest bowiem, że uwielbienie dla serialu „M jak miłość” jest średnio rzecz biorąc domeną kobiet, natomiast deklaracje przeczytania najbardziej znanego dzieła Jamesa Joyce’a (i to w całości) nie zależą tyle od płci, co raczej od wykształcenia. Wydaje się zatem, że w każdej z czterech możliwych grup respondentów charakteryzowanych przez parę (płeć, wykształcenie) uwielbienie dla serialu i fakt przeczytania „Ulissesa” będzie się różnicować w znacznie mniejszym stopniu niż pomiędzy tymi grupami. Jest to właśnie nieformalne sformułowanie tezy, że zmienne X_1 i X_3 są statystycznie⁶ niezależne pod warunkiem zmiennych X_2 i X_4 .

Powyższej sytuacji odpowiada graf \mathcal{G} , który wygląda następująco



⁶Nie będziemy w tej pracy formalizować tego przysłowka. Całkiem nieformalnie można przyjąć, że „statystycznie” znaczy tyle co „w przybliżeniu”.

Rzeczywiście. Brak krawędzi $(1, 3)$ oznacza, że graf ten odpowiada modelowi niezależności $X_1 \perp\!\!\!\perp X_3 | X_2, X_4$. Wszystkie inne krawędzie istnieją, więc model nie reprezentuje żadnych innych niezależności typu $X_i \perp\!\!\!\perp X_j | \text{reszta}$. Graf \mathcal{G} posiada dwie klikki $\{1, 2, 4\}$ i $\{2, 3, 4\}$, a zatem istnieją funkcje nieujemne Ψ_{124} oraz Ψ_{234} takie, że:

$$\mathbb{P}(X_1 = i_1, X_2 = i_2, X_3 = i_3, X_4 = i_4) = \frac{1}{Z} \Psi_{124}(i_1, i_2, i_4) \cdot \Psi_{234}(i_2, i_3, i_4). \quad (2.8)$$

Model graficzny $\{\mathcal{X}, \mathcal{G}, \mathcal{P}_{\mathcal{G}}\}$ w tym wypadku to przestrzeń obserwacji $\mathcal{X} = \{0, 1\}^4$, powyższy graf \mathcal{G} oraz rodzina łącznych rozkładów prawdopodobieństwa, które dają się przedstawić w postaci 2.8.

Jak w ogólnym przypadku pokazaliśmy w poprzednim podrozdziale, dla trzech binarnych zmiennych losowych i modelu $X \perp\!\!\!\perp Y | Z$ mamy geometryczną parametryzację $\mathbb{R}^8 \rightarrow \mathbb{R}^8$ daną wzorem 2.5. Postać równania 2.7 pokazuje, że podobną parametryzację możemy skonstruować dla modeli graficznych. Rozważmy zbiór \mathcal{C} klik grafu \mathcal{G} , niech $d = \sum_{F \in \mathcal{C}} 2^{|F|}$. Wprowadźmy d parametrów t_1, \dots, t_d , które odpowiadają wartościom $\Psi_F(i_F)$ dla wszystkich możliwych $F \in \mathcal{C}$ i wszystkich możliwych i_F . Niech $m = 2^n$. Otrzymujemy wówczas parametryzację jednomianową $\mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^m$:

$$\phi_A \quad : \quad \mathbb{R}^d \rightarrow \mathbb{R}^m, \quad (t_1, \dots, t_d) \mapsto \left(\prod_{i=1}^d t_i^{a_{i1}}, \prod_{i=1}^d t_i^{a_{i2}}, \dots, \prod_{i=1}^d t_i^{a_{im}} \right). \quad (2.9)$$

Aby zobaczyć, ile wynoszą wykładniki a_{ij} , wystarczy dla przykładu rozważyć pierwszą z 2^n współrzędnych w przeciwdziedzinie. Zgodnie z naszą konwencją porządkowania wektorów bazowych przestrzeni $\mathbb{R}^m = \otimes_{i=1}^n \mathbb{R}^2$ w sposób leksykograficzny, pierwsza współrzędna odpowiada sytuacji $(i_1, \dots, i_n) = (0, \dots, 0)$. W tym przypadku równanie 2.7 ma postać

$$\mathbb{P}(X_1 = 0, X_2 = 0, \dots, X_n = 0) = \frac{1}{Z} \prod_{F \in \mathcal{C}} \Psi_F(i_F = (0, \dots, 0)).$$

Wektor $a_1 = [a_{11}, \dots, a_{d1}]$ składa się z $|\mathcal{C}|$ bloków, z których każdy jest długości takiej ile jest wierzchołków w każdej klicie. Blok odpowiadający klicie F odpowiada przestrzeni $\otimes_{i=1}^{|F|} \mathbb{R}^2$. Ustaliliśmy, że wektory bazowe tych przestrzeni również porządkujemy leksykograficznie. Zatem odpowiednie $i_F = (0, \dots, 0)$ odpowiada pierwszej współrzędnej w każdym bloku. Zatem wektor a_1 skonstruowany jest tak, że ma jedynki na pierwszej współrzędnej każdego z bloków i zera poza tym. W podobny sposób odnajdujemy resztę wykładników a_{ij} z parametryzacji 2.9. Zatem macierz A jest $d \times m$ -wymiarową macierzą złożoną z zer i jedynek.

Przykład 2.11 *Rozważmy model graficzny z poprzedniego przykładu. Przestrzeń prób zadana jest tutaj przez zbiór $\{0, 1\}^4$. Mamy zatem $m = 2^4 = 16$ możliwości realizacji czterech binarnych zmiennych losowych. Graf \mathcal{G} ma dwie klikki, co daje 8 możliwych wartości i_F dla $F_1 = \{1, 2, 4\}$ i dla $F_2 = \{2, 3, 4\}$. Mamy $d = 2^3 + 2^3 = 16$ parametrów t_i . Przykładowo*

$$x_{0001} = \text{"}\mathbb{P}(X_1 = X_2 = X_3 = 0, X_4 = 1)\text{"} = t_{001}^{(1,2,4)} t_{001}^{(2,3,4)}.$$

Tutaj na przykład wartości $\Psi_{\{1,2,4\}}(001)$ odpowiada parametr $t_{001}^{(1,2,4)}$ i podobnie dla reszty. Wektor macierzy A odpowiadający wektorowi e_{0001} to $a_2 = [0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0]$. Ogólnie otrzymujemy:

$$x_{ijkl} = t_{ijl}^{(1,2,4)} t_{jkl}^{(2,3,4)}$$

dla wszystkich $(i, j, k, l) \in \{0, 1\}^4$.

Możemy teraz sformułować algebraiczną definicję modelu graficznego.

Definicja 2.12 *Algebraicznym modelem graficznym nazywamy zbiór $\phi_A(\mathbb{R}^d) \cap \Delta_{m-1}$, gdzie ϕ_A dane jest w równaniu 2.9.*

Zgodnie z uwagą 2.6 w rozważaniach geometrycznych będziemy analizować całą sytuację nad ciałem liczb zespolonych. Na mocy teorii przedstawionej w podrozdziale 1.1, parametryzacja 2.9 rozważana jako przekształcenie między przestrzeniami zespolonymi, zadaje rozmaitość semitoryczną. Łatwo zauważyć, że suma elementów w każdej kolumnie macierzy A musi być taka sama (równa liczbie klik w grafie), a zatem w przestrzeni rozpinanej przez jej wiersze leży wektor jedynek, a w konsekwencji A zadaje rzutową rozmaitość semitoryczną. Analizując geometrię modeli graficznych zainteresowani jesteśmy przede wszystkim przekształceniem dualnym do parametryzacji ϕ_A (porównaj równanie 1.1):

$$\Psi_A : \mathbb{C}[x_{i_1 \dots i_n} | i_j \in [d_j]] \rightarrow \mathbb{C}[t_1, t_1^{-1}, \dots, t_d, t_d^{-1}].$$

W tej pracy będziemy koncentrować się na rodzinie modeli bardzo podobnej do modeli graficznych, którą będziemy nazywać rodziną modeli grafów. Model grafu również zdefiniowany jest przez nieskierowany graf \mathcal{G} . Cała konstrukcja modelu przebiega podobnie tylko miast zbioru klik grafu \mathcal{G} w równaniu 2.7, rozważamy odpowiednie produkt po jego krawędziach. Zatem jak poprzednio $m = \prod_{k=1}^n d_k$ jednak $d = \sum_{(i,j) \in E} d_i d_j$.

Przykład 2.13 *W sytuacji w poprzedniego przykładu, zbiór krawędzi grafu E równy jest $\{(1, 2), (2, 3), (2, 4), (3, 4), (4, 1)\}$. Mamy zatem $d = 5 \cdot 4 = 20$ parametrów t_i . Parametryzacja modelu grafu \mathcal{G} zadana jest przez:*

$$x_{ijkl} = t_{ij}^{(1,2)} t_{jk}^{(2,3)} t_{jl}^{(2,4)} t_{kl}^{(3,4)} t_{li}^{(4,1)}$$

dla wszystkich $(i, j, k, l) \in D$.

Oczywiście jeśli graf \mathcal{G} nie zawiera podgrafów w postaci cykli długości trzy, to rodzina modeli graficznych dla \mathcal{G} oraz rodzina modeli grafu \mathcal{G} są sobie równe (mamy te same parametryzacje). Na przykład dla modeli cyklu długości cztery, gdzie zbiór krawędzi to $E = \{(1, 2), (2, 3), (3, 4), (4, 1)\}$ łatwo zauważyć, że zbiór klik tego grafu równy jest zbiorowi krawędzi, a zatem graf definiuje tą samą parametryzację w obu przypadkach.

Rozdział 3

Geometria binarnych modeli grafów

Będziemy stosować notację z podrozdziału 2.4. W tym rozdziale zajmiemy się binarnymi modelami grafów i rozmaitościami, które pojawiają się w ich kontekście. Jak pokazaliśmy wcześniej model i jego parametryzacja zadają macierz A . Macierz ta zadaje wielościan kratowy (patrz rozdział 1). W tym rozdziale zaprezentujemy na początek bardziej geometryczną definicję binarnego modelu grafu. Głównym wynikiem tej pracy jest pokazanie, że w przypadku binarnych modeli grafów bez cykli, wielościan ten jest normalny, a zatem rzutowa rozmaitość semitoryczna Y_A jest normalna. Ważną częścią tego rozdziału będą również rozważania dotyczące wielościanu związanego z binarnym modelem cyklicznym \mathcal{C}_n .

3.1. Model i parametryzacja

Binarny model grafu Jak już zauważyliśmy na końcu poprzedniego rozdziału, konstrukcja modeli grafów przebiega identycznie jak konstrukcja modeli graficznych z tym, że generatorem modelu jest nie zbiór klik grafu lecz zbiór jego krawędzi. W tym podrozdziale definiując w sposób formalny binarny model grafu, będziemy rozpatrywać całą sytuację nad ciałem liczb zespolonych. Będziemy opierać się na notacji z artykułu Develina i Sullivana [DS03]. Sama definicja modelu grafu będzie przebiegała w drugą stronę niż wcześniej. W poprzednim rozdziale to parametryzacja wynikała z postaci grafu, a w konsekwencji mogliśmy otrzymać odpowiedni homomorfizm krat reprezentowany przez macierz A . Tutaj rozpoczniemy od definicji homomorfizmu krat. Z punktu widzenia dalszej analizy kolejność definiowania obiektów nie ma znaczenia, jednak warto pamiętać, że w zastosowaniach w statystyce to parametryzacja jest obiektem pierwotnym.

Niech \mathcal{G} będzie nieskierowanym grafem o zbiorze wierzchołków $V = \{1, 2, \dots, n\}$ oraz zbiorem krawędzi $E(\mathcal{G})$. Przez $\text{Iso}(\mathcal{G})$ będziemy oznaczać zbiór wierzchołków izolowanych grafu \mathcal{G} . Niech $\mathbb{Z}^2 = \text{span}\{e_0, e_1\}$. Będziemy rozważać n kopii kraty \mathbb{Z}^2 , po jednej dla każdego wierzchołka grafu \mathcal{G} . Generatory każdej z tych kopii będziemy dodatkowo etykietować superskryptem z numerem wierzchołka. Dla każdej krawędzi $\{j, k\}$ grafu \mathcal{G} rozważmy homomorfizm krat

$$\pi_{j,k} : \bigotimes_{i=1}^n \mathbb{Z}^2 \rightarrow \mathbb{Z}^2 \otimes \mathbb{Z}^2, \quad e_{i_1}^{(1)} \otimes e_{i_2}^{(2)} \otimes \dots \otimes e_{i_n}^{(n)} \mapsto e_{i_j}^{(j)} \otimes e_{i_k}^{(k)},$$

a dla każdego wierzchołka izolowanego k homomorfizm krat postaci

$$\pi_k : \bigotimes_{i=1}^n \mathbb{Z}^2 \rightarrow \mathbb{Z}^2, \quad e_{i_1}^{(1)} \otimes e_{i_2}^{(2)} \otimes \dots \otimes e_{i_n}^{(n)} \mapsto e_{i_k}^{(k)},$$

gdzie $(i_1, i_2, \dots, i_n) \in \{0, 1\}^n$.

Notacja 3.1 W dalszej części pracy wektory $e_{i_1}^{(1)} \otimes e_{i_2}^{(2)} \otimes \dots \otimes e_{i_n}^{(n)}$ będziemy skrótowo oznaczać przez $e_{i_1 \dots i_n}$, a wektory $e_{i_j}^{(j)} \otimes e_{i_k}^{(k)}$ przez $e_{i_j i_k}^{(j,k)}$. W szczególności zatem

$$\mathbb{Z}^2 \otimes \mathbb{Z}^2 = \text{span}_{\mathbb{Z}}\{e_0 \otimes e_0, e_0 \otimes e_1, e_1 \otimes e_0, e_1 \otimes e_1\} = \text{span}_{\mathbb{Z}}\{e_{00}, e_{01}, e_{10}, e_{11}\}.$$

Definiujemy homomorfizm krat

$$\begin{aligned} A_{\mathcal{G}} : \bigotimes_{i=1}^n \mathbb{Z}^2 &\rightarrow \bigoplus_{\{j,k\} \in E(\mathcal{G})} \mathbb{Z}^2 \otimes \mathbb{Z}^2 \oplus \bigoplus_{\{k\} \in \text{Iso}(\mathcal{G})} \mathbb{Z}^2 \\ v &\mapsto \bigoplus_{\{j,k\} \in E(\mathcal{G})} \pi_{j,k}(v) \oplus \bigoplus_{\{k\} \in \text{Iso}(\mathcal{G})} \pi_k(v). \end{aligned} \quad (3.1)$$

Macierz, która reprezentuje homomorfizm $A_{\mathcal{G}}$ w bazie kraty $\bigotimes_{i=1}^n \mathbb{Z}^2$ równej

$$\{e_{i_1 \dots i_n} : (i_1, \dots, i_n) \in \{0, 1\}^n\}$$

będziemy również oznaczać przez $A_{\mathcal{G}}$. Przy zadanym porządku na zbiorze krawędzi i na zbiorze wierzchołków izolowanych, do pełnej identyfikacji macierzy $A_{\mathcal{G}}$ brakuje uporządkowania elementów bazy krat $\bigotimes_{i=1}^n \mathbb{Z}^2$, $\mathbb{Z}^2 \otimes \mathbb{Z}^2$ i \mathbb{Z}^2 pozwalającego na utożsamienie wektorów bazowych z wektorami jednostkowymi z baz standardowych odpowiednio \mathbb{Z}^{2^n} , \mathbb{Z}^4 i \mathbb{Z}^2 . Wektory bazowe tych krat indeksowane są ciągami zero-jedynkowymi. Porządkować będziemy je w sposób leksykograficzny zakładając $0 >_{\text{lex}} 1$. Jeżeli na przykład założymy, że $n = 4$, to baza kraty $\bigotimes_{i=1}^n \mathbb{Z}^2$ ma szesnaście elementów $0000 >_{\text{lex}} 0001 >_{\text{lex}} 0010 >_{\text{lex}} \dots >_{\text{lex}} 1111$.

Niech wektory $a_1, \dots, a_{2^n} \in \bigoplus_{\{j,k\} \in E(\mathcal{G})} \mathbb{Z}^2 \otimes \mathbb{Z}^2 \oplus \bigoplus_{\{k\} \in \text{Iso}(\mathcal{G})} \mathbb{Z}^2$ stanowią kolejne kolumny macierzy $A_{\mathcal{G}}$. Rozważmy homomorfizm pierścieni wielomianów (porównaj równanie 1.1):

$$\begin{aligned} \psi_{\mathcal{G}} : \mathbb{C}[x_{i_1 \dots i_n} : (i_1, \dots, i_n) \in \{0, 1\}^n] &\rightarrow \mathbb{C}[t_{i_j i_k}^{(j,k)}, t_{i_l}^{(l)} : \{j, k\} \in E(\mathcal{G}), l \in \text{Iso}(\mathcal{G})] \\ p_{i_1 \dots i_n} &\mapsto \prod_{\{j,k\} \in E(\mathcal{G})} t_{i_j i_k}^{(j,k)} \prod_{\{l\} \in \text{Iso}(\mathcal{G})} t_{i_l}^{(l)} = \mathbf{t}^{a_i}, \end{aligned} \quad (3.2)$$

gdzie a_i w jednomianie \mathbf{t}^{a_i} to kolumna $A_{\mathcal{G}}$ odpowiadająca wektorowi $e_{i_1 \dots i_n}$ z bazy $\bigotimes_{i=1}^n \mathbb{Z}^2$. Prawdziwość równości

$$\prod_{\{j,k\} \in E(\mathcal{G})} t_{i_j i_k}^{(j,k)} \prod_{\{l\} \in \text{Iso}(\mathcal{G})} t_{i_l}^{(l)} = \mathbf{t}^{a_i}$$

wynika z samej definicji homomorfizmu krat $A_{\mathcal{G}}$ oraz postaci jego macierzy.

Obiektem naszego zainteresowania będzie binarny model grafu \mathcal{G} , który będziemy w tym rozdziale utożsamiać z rozmaitością generowaną przez ideał $I_{\mathcal{G}} = \ker(\psi_{\mathcal{G}})$, który na mocy definicji z podrozdziału 1.1 jest ideałem torycznym. Z drugiej strony badając rozmaitość generowaną przez ideał $I_{\mathcal{G}}$ możemy, jak pokazują wyniki przedstawione w podrozdziale 1.1, rozważać równoważnie parametryzację jednomianową:

$$\phi_{\mathcal{G}} : \mathbb{R}^d \rightarrow \mathbb{R}^{2^n}, \quad \mathbf{t} = (t_1, \dots, t_d) \mapsto (\mathbf{t}^{a_1}, \mathbf{t}^{a_2}, \dots, \mathbf{t}^{a_{2^n}}), \quad (3.3)$$

gdzie $d = 4 \cdot |E(\mathcal{G})| + 2 \cdot \text{Iso}(\mathcal{G})$ oraz t_1, \dots, t_d stanowią zbiór parametrów odpowiadających wierzchołkom izolowanym i krawędziom grafu \mathcal{G} . Tak jak w rozdziale drugim, algebraiczna parametryzacja tej postaci odpowiada parametryzacji postaci 2.7

Przykład 3.2 Rozważmy graf \mathcal{G} o czterech wierzchołkach, gdzie $E = \{\{1, 2\}, \{2, 3\}\}$ oraz $\text{Iso}(\mathcal{G}) = \{4\}$. Przekształcenie $\psi_{\mathcal{G}}$ idzie z pierścienia wielomianów od szesnastu zmiennych do pierścienia wielomianów od dziesięciu zmiennych

$$\psi_{\mathcal{G}} : \mathbb{C}[x_{ijkl} : (i, j, k, l) \in \{0, 1\}^4] \rightarrow \mathbb{C}[r_{ij}, s_{jk}, t_l]$$

$$x_{ijkl} \mapsto r_{ij} \cdot s_{jk} \cdot t_l.$$

Dla $i = j = k = l = 0$ otrzymujemy $x_{0000} \mapsto r_{00} \cdot s_{00} \cdot t_0$. Pierwsza kolumna macierzy A_G odpowiada obrazowi przy A_G wektora kratowego $e_{0000} \in \mathbb{Z}^{16}$, czyli $e_{00}^{(1,2)} \oplus e_{00}^{(2,3)} \oplus e_0^{(4)}$, co we współrzędnych kraty \mathbb{Z}^{10} daje wektor $[1, 0, 0, 0; 1, 0, 0, 0; 1, 0]$, gdzie średnikami oddzielono bloki związane z poszczególnymi krawędziami i wierzchołkami izolowanymi. Istotnie zatem

$$\mathbf{t}^{a_1} = r_{00}^1 \cdot r_{01}^0 \cdot r_{10}^0 \cdot r_{11}^0 \cdot s_{00}^1 \cdot s_{01}^0 \cdot s_{10}^0 \cdot s_{11}^0 \cdot t_0^1 \cdot t_1^0 = r_{00} \cdot s_{00} \cdot t_0.$$

Sprawdzenie w podobny sposób dla pozostałych wektorów z bazy kraty \mathbb{Z}^{16} pokazuje, że kolumny macierzy A_G indukują parametryzację jednomianową jak w równaniu 3.3.

W rozważaniach geometrycznych niezwykle istotny okaże się wielościan będący kombinacją wypukłą kolumn macierzy A_G . Będziemy go w ogólności oznaczać jako P_G . Związek tego wielościanu z rozmaitością semitoryczną definiowaną przez A_G jest opisywany bardziej szczegółowo w rozdziale pierwszym. Zauważmy, że z samej definicji homomorfizmu A_G , suma współrzędnych w każdej kolumnie A_G wynosi tyle samo i jest równa $|E(\mathcal{G})| + |\text{Iso}(\mathcal{G})|$. W związku z tym w szczególności w przestrzeni liniowej rozpinanej przez wiersze macierzy A_G leży wektor jedynek. W związku z tym każdy dwumian w $\ker_{\mathbb{Z}} A_G$ jest jednorodny (patrz podrozdział 1.1, a A_G zadaje rzutową rozmaitość semitoryczną. W tej sytuacji będziemy również badać półgrupę S_{P_G} oraz wielościenne pierścień półgrupowy $\mathbb{C}[S_{P_G}]$ (porównaj podrozdział 1.1).

Binarny model cykliczny Zbiorem generatorów modelu grafu cyklicznego jest zbiór krawędzi grafu cyklicznego \mathcal{C}_n

$$E(\mathcal{C}_n) = \left\{ \{1, 2\}, \{2, 3\}, \dots, \{n, 1\} \right\}.$$

Przez A_n będziemy oznaczać macierz będącą macierzą homomorfizmu krat definiowanego przez graf \mathcal{C}_n analogicznie jak w równaniu 3.1. Rozważania wcześniejszego akapitu pokazują, że $A_n = (a_{ij})$ jest macierzą złożoną z zer i jedynek wymiaru $4n \times 2^n$. Homomorfizm krat A_n zdefiniowany jest w tym przypadku jako:

$$e_{i_1 \dots i_n} \mapsto e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \dots \oplus e_{i_{n-1} i_n}^{(n-1)} \oplus e_{i_n i_1}^{(n)}. \quad (3.4)$$

Czyli stosujemy konwencję numerowania elementów $E(\mathcal{C}_n)$ tak, że krawędź $\{i, i+1\}$ ma numer i oraz przyjmujemy, że $n+1 \equiv 1$. Jak łatwo zauważyć, w tym przypadku suma elementów w każdej kolumnie macierzy A_n wynosi n .

Uliniowanie kraty $\bigoplus_{e \in \mathcal{C}_n} \mathbb{Z}^2 \otimes \mathbb{Z}^2 = \mathbb{Z}^{4n}$, czyli przestrzeń $\bigoplus_{i=1}^n \mathbb{R}^2 \otimes \mathbb{R}^2 = \mathbb{R}^{4n}$ będziemy odtąd oznaczać przez V_n . Bazę standardową V_n będziemy zapisywać jako

$$\{e_{00}^{(1)}, e_{01}^{(1)}, e_{10}^{(1)}, e_{11}^{(1)}; e_{00}^{(2)}, e_{01}^{(2)}, e_{10}^{(2)}, e_{11}^{(2)}; \dots; e_{00}^{(n)}, e_{01}^{(n)}, e_{10}^{(n)}, e_{11}^{(n)}\},$$

a jej współrzędne będziemy zapisywać jako

$$(x_{00}^{(1)}, x_{01}^{(1)}, x_{10}^{(1)}, x_{11}^{(1)}; x_{00}^{(2)}, x_{01}^{(2)}, x_{10}^{(2)}, x_{11}^{(2)}; \dots; x_{00}^{(n)}, x_{01}^{(n)}, x_{10}^{(n)}, x_{11}^{(n)}),$$

gdzie średnikami oddzielamy od siebie każdy z n bloków przestrzeni V_n , które stanowią wspomniane wcześniej uliniowanie krat $\mathbb{Z}^2 \otimes \mathbb{Z}^2$.

Jak wynika z równania 3.4, macierz A_n na przykład dla $n = 4$ przedstawia się jak poniżej:

$$A_4 = \begin{matrix} x_{00}^{(1)} \\ x_{01}^{(1)} \\ x_{10}^{(1)} \\ x_{11}^{(1)} \\ x_{00}^{(2)} \\ x_{01}^{(2)} \\ x_{10}^{(2)} \\ x_{11}^{(2)} \\ x_{00}^{(3)} \\ x_{01}^{(3)} \\ x_{10}^{(3)} \\ x_{11}^{(3)} \\ x_{00}^{(4)} \\ x_{01}^{(4)} \\ x_{10}^{(4)} \\ x_{11}^{(4)} \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

gdzie w wierszach dopisano dodatkowo współrzędne przestrzeni V_4 , którym odpowiadają.

Tak jak to zostało przedstawione równaniem 3.2, przekształcenie A_n indukuje również homomorfizm pierścieni wielomianów:

$$\begin{aligned} \psi_n &: \mathbb{C}[x_{i_1 \dots i_n} | i_k \in \{0, 1\}] \rightarrow \mathbb{C}[t_{i_k i_{k+1}}^{(k)} | k = 1, \dots, n; i_k \in \{0, 1\}] \\ x_{i_1 \dots i_n} &\mapsto t_{i_1 i_2}^{(1)} \cdot t_{i_2 i_3}^{(2)} \cdot \dots \cdot t_{i_{n-1} i_n}^{(n-1)} \cdot t_{i_n i_1}^{(n)} \end{aligned} \quad (3.5)$$

Skoro dla wszystkich kolumn macierzy A_n zachodzi $x_{00}^{(i)} + x_{01}^{(i)} + x_{10}^{(i)} + x_{11}^{(i)} = 1$ dla wszystkich $i = 1, \dots, n$, to możemy rozważać rzutową wersję parametryzacji binarnego modelu cyklicznego, która zdefiniowana jest jako:

$$\begin{aligned} \phi_n &: \prod_{i=1}^n \mathbb{P}^3 \rightarrow \mathbb{P}^{2^n - 1} \\ (t_{00}^{(1)}, t_{01}^{(1)}, t_{10}^{(1)}, t_{11}^{(1)}; \dots; t_{00}^{(n)}, t_{01}^{(n)}, t_{10}^{(n)}, t_{11}^{(n)}) &\mapsto \left(\prod_{k=1}^n t_{i_k i_{k+1}}^{(k)} \right)_{(i_1, \dots, i_n) \in \{0, 1\}^n}, \end{aligned} \quad (3.6)$$

gdzie $[t_{00}^{(i)}, t_{01}^{(i)}, t_{10}^{(i)}, t_{11}^{(i)}]$ są współrzędnymi jednorodnymi na każdej \mathbb{P}^3 z powyższego produktu. Wówczas zgodnie z definicją 2.5 binarny model cykliczny jest domknięciem obrazu tego przekształcenia.

Binarny model liniowy Grafy liniowe to grafy, które uzyskuje się z grafów cyklicznych po usunięciu jednej krawędzi, a zatem zbiór krawędzi to $E = \{\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}\}$. Binarny model liniowy dla grafu o n wierzchołkach będziemy oznaczać przez \mathcal{L}_n . Wielościan R_n dla tego modelu zadany jest jako zbiór

$$\text{conv} \left\{ e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \dots \oplus e_{i_{n-1} i_n}^{(n-1)} \mid (i_1, i_2, \dots, i_n) \in \{0, 1\}^n \right\} \subset V_{n-1}, \quad (3.7)$$

gdzie jak ustaliliśmy poprzednio $V_{n-1} = \mathbb{R}^{4(n-1)}$. Postać tego wielościanu uzyskujemy w sposób analogiczny jak w przypadku binarnych modeli cyklicznych. Najprostszy przykład dla $n = 3$ podany jest w dodatku B.1. Współrzędne przestrzeni V_{n-1} będziemy zapisywać jako

$$(x_{00}^{(1)}, x_{01}^{(1)}, x_{10}^{(1)}, x_{11}^{(1)}; \dots; x_{00}^{(n-1)}, x_{01}^{(n-1)}, x_{10}^{(n-1)}, x_{11}^{(n-1)}),$$

gdzie jak poprzednio, każdy z $n-1$ bloków przestrzeni V_{n-1} będziemy oddzielać średnikami.

3.2. Normalność wielościanów niektórych modeli grafów

Podstawowym obiektem naszych zainteresowań w tym podrozdziale stanie się wielościan R_n indukowany z parametryzacji binarnego modelu liniowego oraz wielościan P_G dla dowolnego modelu grafu G bez cykli. Z samej definicji wielościanów dla binarnych modeli grafów wynika, że R_n i P_G są 0/1-wielościanami, tzn. zbiór ich wierzchołków stanowi podzbiór wierzchołków standardowej kostki. Wierzchołki 0/1-wielościanów są jego jedynymi punktami kratowymi. Pokażemy, że P_G (oraz R_n jako szczególny przypadek) mają pokrycie unimodularne względem kraty generowanej przez swoje wierzchołki. Istnienie pokrycia unimodularnego zapewni, że rzutowa różnorodność semitoryczna definiowana przez te modele jest normalna, co wynika z wniosku 1.12.

3.2.1. Unimodularność dla modeli grafów liniowych

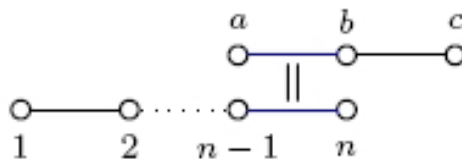
Dowód istnienia pokrycia unimodularnego dla wielościanów związanych z binarnymi modelami grafów liniowych \mathcal{L}_n będzie podobny jak w [BW06] (tam obiektem zainteresowania były modele drzew filogenetycznych). Będzie on polegał na sklejanu odpowiednich grafów, którym odpowiadają wielościany normalne. Jak się okaże, sklejenie dwóch grafów wzdłuż krawędzi odpowiada rozważaniu produktu włóknistego (patrz definicja A.3) dwóch wielościanów kratowych nad simpleksem.

Zauważmy, że z definicji wielościanu R_n (równanie 3.7) wynika, że dla każdego $i = 1, \dots, n-1$ zachodzi $x_{00}^{(i)} + x_{01}^{(i)} + x_{10}^{(i)} + x_{11}^{(i)} = 1$, czyli w każdym z $n-1$ bloków znajomość trzech współrzędnych determinuje czwartą. Motywuje to następującą konstrukcję.

Konstrukcja 3.3 Niech $R_n \subset (M_1)_{\mathbb{R}}$ i $R_3 \subset (M_2)_{\mathbb{R}}$ będą dwoma wielościanami kratowymi związanymi odpowiednio z modelem \mathcal{L}_n oraz \mathcal{L}_3 dla $n \geq 2$. Rozważmy dwa homomorfizmy krat $\ell_i : M_i \rightarrow \mathbb{Z}^3$ takie, że:

- $(\ell_1)_{\mathbb{R}}(R_n) = \Delta_3$, gdzie Δ_3 to standardowy simpleks 3-wymiarowy, jest rzutowaniem na ostatnie trzy współrzędne $[x_{01}^{(n-1)}, x_{10}^{(n-1)}, x_{11}^{(n-1)}]$.
- $(\ell_2)_{\mathbb{R}}(R_3) = \Delta_3$ jest rzutowaniem na współrzędne: drugą, trzecią i czwartą $[x_{01}^{(a)}, x_{10}^{(a)}, x_{11}^{(a)}]$.

Na poziomie grafów właściwy produkt włóknisty nad Δ_3 odpowiada poniższemu sklejeniu:



Twierdzimy, że wielościan $R_n \times_{\ell_1 \times \ell_2} R_3$ będący produktem włóknistym nad Δ_3 jest izomorficzny w sensie izomorfizmu krat z wielościanem $R_{n+1} \subset M_{\mathbb{R}}$. Powyższe sklejenie pozwala zatem skonstruować wielościan odpowiadający modelowi \mathcal{L}_{n+1} .

Dowód: Na mocy definicji wielościanu R_n wielościan $R_n \times R_3$ ma wierzchołki postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \dots \oplus e_{i_{n-1} i_n}^{(n-1)} \oplus e_{j_a j_b}^{(a)} \oplus e_{j_b j_c}^{(b)},$$

gdzie $e_{ij}^{(k)}$ dla $k = 1, \dots, n-1$ to wektory przestrzeni V_{n-1} , a $e_{ij}^{(k)}$ dla $k = a, b$ to wektory przestrzeni V_2 oraz $(i_1, \dots, i_n) \in \{0, 1\}^n$, $(j_a, j_b, j_c) \in \{0, 1\}^3$. W przestrzeni $\ker(\ell_1 - \ell_2)$ znajdują się te spośród nich, które są postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \dots \oplus e_{i_{n-1} i_n}^{(n-1)} \oplus e_{i_{n-1} i_n}^{(a)} \oplus e_{i_n j}^{(b)}, \quad (3.8)$$

gdyż tylko dla tych wierzchołków zachodzą zgodności na odpowiednich współrzędnych. Wynika to z tego, że równości

$$x_{01}^{(a)} = x_{01}^{(n-1)}, x_{01}^{(a)} = x_{01}^{(n-1)}, x_{01}^{(a)} = x_{01}^{(n-1)}$$

na mocy przedstawionej wcześniej równości $x_{00}^{(i)} + x_{01}^{(i)} + x_{10}^{(i)} + x_{11}^{(i)} = 1$ implikują również, że $x_{00}^{(a)} = x_{00}^{(n-1)}$. Wierzchołki w równaniu 3.8 to jedyne spośród wierzchołków $R_n \times R_3$, które wszystkie te cztery równości spełniają jednocześnie. Wielościan rozpinany przez te wyróżnione wierzchołki to $R_n \times_{\ell_1 \times \ell_2} R_3$. Istotnie, na mocy uwagi A.7 produkt włóknisty nad simpleksem jest wielościanem kratowym. Ale skoro $R_n \times R_3$ jest 0/1-wielościanem, to jedynymi wierzchołkami wielościanu $R_n \times_{\ell_1 \times \ell_2} R_3$ mogą być wierzchołki z równania 3.8.

Łatwo zauważyć, że wielościany $R_n \times_{\ell_1 \times \ell_2} R_3$ oraz $R_{n+1} \subset M_{\mathbb{R}}$ są izomorficzne (jako wielościany kratowe). Jeśli utożsamimy superskrypty b i n , to przekształcenie będące homomorfizmem krat $f : M_1 \times_{\ell_1 \times \ell_2} M_2 \rightarrow M$ definiujemy jako:

$$f(e_{ij}^{(k)}) = e_{ij}^{(k)} \text{ dla } k \neq a; \quad f(e_{ij}^{(a)}) = 0$$

dla wszystkich par (i, j) . Przekształcenie odwrotne $g : M \rightarrow M_1 \times_{\ell_1 \times \ell_2} M_2$ definiują równości:

$$g(e_{ij}^{(k)}) = e_{ij}^{(k)} \text{ dla } k \neq n-1; \quad g(e_{ij}^{(n-1)}) = e_{ij}^{(n-1)} + e_{ij}^{(a)}.$$

Co dowodzi poprawności konstrukcji. \square

Następujący lemat jest natychmiastowy na mocy powyższej konstrukcji oraz wniosku A.8.

Lemat 3.4 (Unimodularność dla modelu liniowego) *Wielościan R_n związany z binarnym modelem liniowym \mathcal{L}_n dla $n \geq 0$ ma pokrycie unimodularne względem kraty generowanej przez jego wierzchołki.*

Dowód: Argument będzie miał charakter indukcyjny. Przypadki $n = 2, 3$ umiemy wyliczyć bezpośrednio (korzystaliśmy z programu `Polymake` oraz skryptu zamieszczonego w dodatku B.3). Wielościan dla modelu \mathcal{L}_{n+1} dostajemy skleając ostatnią krawędź grafu dla modelu \mathcal{L}_n z pierwszą krawędzią grafu modelu \mathcal{L}_3 , jak w konstrukcji 3.3. Na mocy wniosku A.8 oraz założenia indukcyjnego dostajemy, że wielościan ten ma pokrycie unimodularne względem kraty generowanej przez jego wierzchołki. Krata ta jest równoważna z kratą będącą odpowiednim produktem włóknistym krat z konstrukcji 3.3. \square

Wniosek 3.5 *Wymiar wielościanu R_n dla binarnego modelu liniowego \mathcal{L}_n wynosi $2n - 1$.*

Dowód: Bezpośrednie obliczenia pokazują, że $\dim R_2 = 3$ (R_2 jest izomorficzne z simpleksem) oraz $\dim R_3 = 5$. Skoro $\dim(R_n \times R_3) = \dim(R_n) + \dim(R_3)$, to dowód lematu 3.4 pokazuje, że $\dim(R_{n+1}) = \dim(R_n \times R_3) - 3$, a zatem

$$\dim(R_{n+1}) - \dim(R_n) = 2, \quad \dim(R_3) = 5.$$

Ostatecznie $\dim(R_n) = 2n - 1$. \square

Lemat 3.6 (Postać wielościanu dla modelu \mathcal{L}_n) *Wielościan R_n dla modelu \mathcal{L}_n jest opisywany przez $2n - 3$ równości*

$$x_{00}^{(i)} + x_{10}^{(i)} = x_{00}^{(i+1)} + x_{01}^{(i+1)} \quad \text{dla } i = 1, \dots, n - 2$$

$$x_{00}^{(i)} + x_{10}^{(i)} + x_{00}^{(i+1)} + x_{01}^{(i+1)} = 1 \quad \text{dla } i = 1, \dots, n - 1$$

oraz $4(n - 1)$ nierówności $x_{ij}^{(k)} \geq 0$ dla $k = 1, \dots, n - 1$. Przy czym pierwszy zestaw $n - 1$ równości można zamienić na równoważny

$$x_{01}^{(i)} + x_{11}^{(i)} = x_{10}^{(i+1)} + x_{11}^{(i+1)} \quad \text{dla } i = 1, \dots, n - 2.$$

Dowód: Mamy $2n - 3$ niezależnych (żadna nie wynika z innych) równości. Łatwo sprawdzić, że muszą być one wszystkie spełnione dla wierzchołków wielościanu R_n z samej jego definicji (równanie 3.7). Kompletność zbioru równości wynika z tego, że $\dim(R_n) = 2n - 1$ (wniosek 3.5). Istotnie, R_n leży w przestrzeni V_{n-1} oraz

$$\dim V_{n-1} - (2n - 3) = 2n - 1,$$

a zatem zachodzi równość ko/wymiarów. Inny możliwy argument pojawi się poniżej przy okazji badania nierówności.

Jeśli chodzi bowiem o postać nierówności, to bezpośrednie obliczenia pokazały, że w przypadku $n = 4$ rzeczywiście nierówności tego typu odpowiadają wszystkim ścianom maksymalnego wymiaru wielościanu R_4 . Przypuśćmy, że postać nierówności jest żądana dla wielościanu R_n . Ale skoro R_n opisywany jest przez nierówności $x_{ij}^{(k)} \geq 0$ dla $k = 1, \dots, n - 1$ oraz R_3 opisywany jest przez nierówności $y_{ij}^{(1)} \geq 0, y_{ij}^{(2)} \geq 0$, to R_{n+1} opisywany jest przez nierówności $x_{ij}^{(k)} \geq 0, y_{ij}^{(2)} \geq 0$ dla $k = 1, \dots, n - 1$, co wynika z konstrukcji 3.3. Oznaczając $y_{ij}^{(2)}$ jako $x_{ij}^{(n)}$ na mocy indukcji dostajemy tezę.

Na koniec zauważmy, że dla każdego i równość $x_{00}^{(i)} + x_{10}^{(i)} = x_{00}^{(i+1)} + x_{01}^{(i+1)}$ jest równoważna z $x_{01}^{(i)} + x_{11}^{(i)} = x_{10}^{(i+1)} + x_{11}^{(i+1)}$ przez wykorzystanie po obu stronach równania typu $x_{00}^{(i)} + x_{10}^{(i)} + x_{00}^{(i+1)} + x_{01}^{(i+1)} = 1$. \square

Wniosek 3.7 *W przypadku punktów $x \in R_n$ współrzędne $\{x_{10}^{(i)}, x_{11}^{(i)}\}_{i=1}^{n-1}$ wraz z $x_{01}^{(n-1)}$ wyznaczają jednoznacznie pozostałe współrzędne punktu x w przestrzeni V_{n-1} .*

Dowód: Skoro znamy współrzędne $\{x_{10}^{(i)}, x_{11}^{(i)}\}_{i=1}^{n-1}$ oraz $x_{01}^{(n-1)}$, to oczywiście znamy również $x_{00}^{(n-1)}$ (w każdym bloku suma współrzędnych równa jeden). Ale mamy dwie równoważne równości (patrz lemat 3.6)

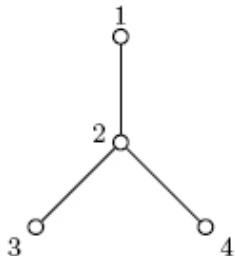
$$x_{00}^{(n-2)} + x_{10}^{(n-2)} = x_{00}^{(n-1)} + x_{01}^{(n-1)}$$

$$x_{01}^{(n-2)} + x_{11}^{(n-2)} = x_{10}^{(n-1)} + x_{11}^{(n-1)},$$

z których wyliczamy $x_{00}^{(n-2)}$ i $x_{01}^{(n-2)}$. Procedurę powtarzamy aż wyznaczymy wszystkie pozostałe współrzędne. Warto zauważyć, że skoro $\dim R_n = 2n - 1$ to liczby współrzędnych w treści wniosku nie da się zredukować. \square

3.2.2. Unimodularność dla modeli grafów bez cykli

W tym paragrafie stosując wyniki uzyskane dla binarnych modeli liniowych, pokażemy, że w istocie dla wszystkich modeli grafów \mathcal{G} bez cykli indukowany wielościan $P_{\mathcal{G}}$ musi być normalny. Pokażemy, że konstrukcja 3.3 daje się w sposób bezpośredni uogólnić. Aby zobrazować tę ideę rozważmy na przykład proste drzewo o czterech wierzchołkach, jak na poniższym rysunku.



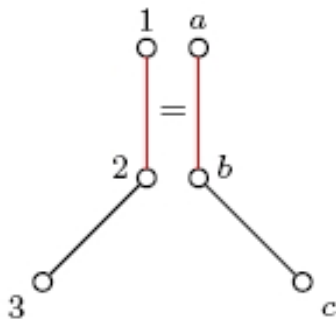
Z ogólnej konstrukcji wielościanów binarnych modeli grafów wynika (porównaj z definicją homomorfizmu 3.1), że wielościan T_4 dla powyższego grafu zadany jest jako zbiór

$$\text{conv}\{e_{i_1 i_2}^{(1,2)} \oplus e_{i_2 i_3}^{(2,3)} \oplus e_{i_2 i_4}^{(2,4)} \mid (i_1, i_2, i_3, i_4) \in \{0, 1\}^4\} \subset V_3. \quad (3.9)$$

Konstrukcja 3.8 Niech $R_3 \subset (M_1)_{\mathbb{R}}$ i $R_3 \subset (M_2)_{\mathbb{R}}$ będą dwoma wielościanami kratowymi związanymi z modelami \mathcal{L}_3 . Rozważmy dwa homomorfizmy krat $\ell_i : M_i \rightarrow \mathbb{Z}^3$ takie, że:

- $(\ell_1)_{\mathbb{R}}(R_3) = \Delta_3$, jest rzutowaniem na współrzędne $[x_{01}^{(1)}, x_{10}^{(1)}, x_{11}^{(1)}]$.
- $(\ell_2)_{\mathbb{R}}(R_3) = \Delta_3$ jest rzutowaniem na współrzędne $[x_{01}^{(a)}, x_{10}^{(a)}, x_{11}^{(a)}]$.

Na poziomie grafów właściwy produkt włóknisty nad Δ_3 odpowiada poniższemu sklejeniu:



Twierdzimy, że wielościan $R_3 \times_{\ell_1 \times \ell_2} R_3$ będący produktem włóknistym nad Δ_3 jest izomorficzny w sensie izomorfizmu krat z wielościanem $T_4 \subset M_{\mathbb{R}}$.

Dowód: Wielościan $R_3 \times R_3$ ma wierzchołki postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus e_{i_a i_b}^{(a)} \oplus e_{i_b i_c}^{(b)}.$$

W podprzestrzeni $\ker(\ell_1 - \ell_2)$ znajdują się te spośród nich, które są postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus e_{i_1 i_2}^{(a)} \oplus e_{i_2 j}^{(b)}, \quad (3.10)$$

gdź tylko dla tych wierzchołków zachodzą zgodności na odpowiednich współrzędnych. Wynika to z tego, że równości

$$x_{01}^{(1)} = x_{01}^{(a)}, x_{10}^{(1)} = x_{10}^{(a)}, x_{01}^{(1)} = x_{01}^{(a)}$$

implikują również, że $x_{00}^{(1)} = x_{00}^{(a)}$. Wierzchołki w powyższym równaniu to jedyne spośród wierzchołków wielościanu $R_3 \times R_3$, które wszystkie te cztery równości spełniają jednocześnie. Wielościan rozpinany przez te wyróżnione wierzchołki to $R_{3 \ell_1 \times \ell_2} R_3$. Istotnie, na mocy uwagi A.7 produkt włóknisty nad simpleksem jest wielościanem kratowym. Ale skoro $R_3 \times R_3$ jest 0/1-wielościanem, to jedynymi wierzchołkami wielościanu $R_{3 \ell_1 \times \ell_2} R_3$ mogą być wierzchołki wymienione powyżej.

Zauważmy, że wielościany $R_{3 \ell_1 \times \ell_2} R_3$ oraz T_4 są izomorficzne (jako wielościany kratowe). Jeśli utożsamimy superskrypty (1) z (1, 2); (2) z (2, 3) oraz (b) z (2, 4), to przekształcenie $f : M_{1 \ell_1 \times \ell_2} M_2 \rightarrow M$ zadane jest jako homomorfizm krat spełniający

$$f(e_{ij}^{(k)}) = e_{ij}^{(k)} \text{ dla } k \neq a; \quad f(e_{ij}^{(a)}) = 0$$

dla wszystkich par (i, j) . Przekształcenie odwrotne $g : M \rightarrow M_{1 \ell_1 \times \ell_2} M_2$ zdefiniowane jest jako:

$$g(e_{ij}^{(k)}) = e_{ij}^{(k)} \text{ dla } k \neq 1; \quad g(e_{ij}^{(1)}) = e_{ij}^{(1)} + e_{ij}^{(a)}.$$

Co dowodzi poprawności konstrukcji. \square

Rozumowanie podobne, jak w dowodzie lematu 3.4 zapewni, że wielościan T_4 ma pokrycie unimodularne w kracie generowanej przez swoje wierzchołki. Warto jeszcze zauważyć, że w istocie wielościany T_4 i R_4 nie są izomorficzne jako wielościany kratowe. W szczególności obliczenia w Polymake pokazują, że choć oba mają dwanaście ścian maksymalnego wymiaru, to różnią się liczbą innych ścian. Liczby ścian kolejnych wymiarów (od zera do sześciu) w przypadku wielościanu T_4 to [16, 88, 204, 242, 160, 60, 12], a w przypadku wielościanu R_4 [16, 80, 180, 216, 148, 58, 12]. Okazuje się zatem, że przy sklejaniu grafów istotne jest jakie wierzchołki ze sobą utożsamiamy, co ustala orientację sklejenia. Dlatego choć zarówno R_4 jak i T_4 uzyskujemy jako produkty włókniste związane ze sklejeniem dwóch grafów R_3 wzdłuż pojedynczej krawędzi, to różna orientacja sklejenia zapewnia, że wielościany te nie są sobie równe.

W istocie przykład z konstrukcją dla wielościanu T_4 pokazał, że sklejjąc ze sobą wzdłuż pojedynczej krawędzi albo grafy liniowe albo grafy uzyskane ze sklejenia wzdłuż pojedynczej krawędzi grafów liniowych i tak dalej, możemy uzyskać dowolny graf spójny bez cykli. Dla każdej z tych procedur sklejenia można pokazać, że na wyjściu dostaniemy binarny model grafu \mathcal{G} , którego wielościan $P_{\mathcal{G}}$ ma pokrycie unimodularne.

Twierdzenie 3.9 *Każde drzewo (spójny graf bez cykli) zadaje binarny model grafu, który definiuje rozmaitość, która jest toryczna.*

Dowód: Załóżmy, że $\mathcal{G} = (V, E)$ jest grafem spójnym bez cykli. Własności drzew zapewniają (patrz [Wil04], rozdział 4, twierdzenie 9.1), że każde dwa wierzchołki ze zbioru V łączy dokładnie jedna droga. Motywuje to poniższą procedurę rekurencyjną, która zapewnia, że model grafu \mathcal{G} odpowiada normalnemu wielościanowi kratowemu, a w konsekwencji (patrz podrozdział 1.2.3) definiuje rozmaitość, która jest toryczna.

1. Niech v_0, w_0 będą parą najodleglejszych wierzchołków w \mathcal{G} (jeśli takich par jest więcej to bierzemy jedną z nich), czyli takich, że droga pomiędzy tymi wierzchołkami jest najdłuższa spośród dróg w \mathcal{G} . Niech $\mathcal{G}_0 = (V_0, E_0)$ będzie podgrafem \mathcal{G} , który łączy v_0 i w_0 (minimalny podgraf spójny, który zawiera v_0 i w_0). \mathcal{G}_0 jest grafem liniowym, a zatem zadaje wielościan normalny. Połóżmy $i = 0$.

2. Niech $u \in V \setminus V_i$ będzie wierzchołkiem \mathcal{G} takim, że w \mathcal{G} istnieje krawędź łącząca go z jednym z wierzchołków \mathcal{G}_i . To, że takie u musi istnieć, wynika ze spójności grafu \mathcal{G} . Co więcej dla każdego takiego u wierzchołek z \mathcal{G}_i , z którym łączy go krawędź jest już zadany jednoznacznie – oznaczmy go v_1 . Niech \mathcal{L}_3 będzie grafem liniowym o trzech wierzchołkach $\{u, w_1, w_2\}$ i krawędziach $\{\{u, w_1\}, \{w_1, w_2\}\}$. Niech $e = (v_1, v_2) \in E_i$ będzie krawędzią \mathcal{G}_i zawierającą wierzchołek v_1 . Graf \mathcal{G}_{i+1} tworzymy poprzez sklejenie grafów \mathcal{G}_i oraz grafu \mathcal{L}_3 wzdłuż krawędzi $\{w_1, w_2\}$ i $\{v_1, v_2\}$ i z utożsamieniem $w_1 \equiv v_1$ oraz $w_2 \equiv v_2$, które zadaje orientację sklejenia.
3. Połóżmy $i = i + 1$. Jeśli $V_i \neq V$ to wracamy do kroku 2. W przeciwnym przypadku procedurę kończymy.

Procedurę powtarzamy rekurencyjnie aż zbiór $V_i = V$. Własność drzewa zapewnia, że dostaniemy odpowiedni wynik, gdyż za każdym razem wybierając wierzchołek u dodawana krawędź jest zadana jednoznacznie.

Z konstrukcji wynika, że w każdym kroku odpowiedniemu sklejeniu grafów odpowiada produkt włóknisty nad simpleksem dwóch normalnych wielościanów kratowych. Na mocy wniosku A.8 modelowi grafu \mathcal{G} odpowiada normalny wielościan kratowy. \square

Lemat 3.10 *Niech \mathcal{G} będzie dowolnym grafem nieskierowanym, który składa się z dwóch rozłącznych składowych \mathcal{G}_1 i \mathcal{G}_2 . Wówczas wielościan P związany z modelem grafu \mathcal{G} jest produktem wielościanów związanych z modelami grafów \mathcal{G}_1 i \mathcal{G}_2 .*

Dowód: Bez straty ogólności możemy założyć, że obie składowe mają więcej niż jeden wierzchołek. W przeciwnym razie suma prosta poniżej będzie wyglądać odrobinę inaczej, ale nie zmieni to rozumowania. Homomorfizm krat $A_{\mathcal{G}}$ związany z modelem grafu, który zdefiniowany jest równaniem 3.1 można w tym przypadku równoważnie zapisać:

$$A_{\mathcal{G}} : \bigotimes_{i=1}^n \mathbb{Z}^2 \rightarrow \bigoplus_{e \in E(\mathcal{G})} \mathbb{Z}^2 \otimes \mathbb{Z}^2 = \bigoplus_{e \in E(\mathcal{G}_1)} \mathbb{Z}^2 \otimes \mathbb{Z}^2 \oplus \bigoplus_{e \in E(\mathcal{G}_2)} \mathbb{Z}^2 \otimes \mathbb{Z}^2$$

$$e_{i_1 \dots i_n} \mapsto \bigoplus_{\{j,k\} \in E(\mathcal{G}_1)} e_{ij i_k}^{(j,k)} \oplus \bigoplus_{\{j,k\} \in E(\mathcal{G}_2)} e_{ij i_k}^{(j,k)},$$

gdzie $E(\mathcal{G}_i)$ to zbiór krawędzi składowej \mathcal{G}_i . Zauważmy, że jeśli oznaczymy

$$P_1 = \text{conv}\{\bigoplus_{\{j,k\} \in E(\mathcal{G}_1)} e_{ij i_k}^{(j,k)} : (i_1, \dots, i_n) \in \{0, 1\}^n\}$$

$$P_2 = \text{conv}\{\bigoplus_{\{j,k\} \in E(\mathcal{G}_2)} e_{ij i_k}^{(j,k)} : (i_1, \dots, i_n) \in \{0, 1\}^n\}$$

to wielościan P jest równy wielościanowi $P_1 \times P_2$. \square

Wniosek 3.11 *Każdy las (graf bez cykli) zadaje binarny model grafu, który definiuje rozmiarowość toryczną.*

Dowód: Niech \mathcal{G} będzie grafem bez cykli. Jeśli \mathcal{G} jest spójny, to mamy tezę na mocy twierdzenia 3.9. Możemy zatem założyć, że \mathcal{G} nie jest spójny. Przez $\mathcal{G}_1, \dots, \mathcal{G}_k$ oznaczmy jego spójne składowe. Na mocy twierdzenia 3.9, każdej składowej odpowiada normalny wielościan kratowy. Na mocy lematu 3.10 każdej parze spójnych składowych odpowiada wielościan, który jest odpowiednim produktem dwóch wielościanów normalnych. Wniosek A.8 mówi, że wielościan ten jest również normalny. Sekwencyjne stosowanie lematu 3.10 oraz wniosku A.8 zapewni prawdziwość tezy. \square

3.2.3. Kwestia unimodularności dla modeli cyklicznych

Grafy bez cykli nie wyczerpują listy grafów definiujących binarne modele grafów, dla których wielościany P_G są normalne. Dla wszystkich n , dla których obliczenia były możliwe ($n = 3, 4, 5, 6, 7$), okazało się, że binarny model cykliczny C_n zadaje wielościan, który posiada nie tylko pokrycie unimodularne, ale nawet triangulację za pomocą simpleksów unimodularnych. Dlatego oczekujemy, że wielościan P_n jest normalny dla dowolnego n . Sturmfels i Sullivant [SS06] przedstawili ostatnio hipotezę, że rozmaitość związana z modelem grafu jest normalna wtedy i tylko wtedy, gdy graf nie posiada podgrafów pełnych o czterech wierzchołkach (podgrafów o czterech wierzchołkach takich, że każde dwa łączą krawędź).

Wielościan P_n dla grafu cyklicznego zadany jest jako zbiór

$$\text{conv}\{e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \cdots \oplus e_{i_{n-1} i_n}^{(n-1)} \oplus e_{i_n i_1}^{(n)} \mid (i_1, i_2, \dots, i_n) \in \{0, 1\}^n\} \subset V_n. \quad (3.11)$$

Najprostszy przykład dla $n = 3$ podany jest w dodatku B.1. Jak w poprzednim paragrafie współrzędne przestrzeni V_n będziemy zapisywać jako

$$x = (x_{00}^{(1)}, x_{01}^{(1)}, x_{10}^{(1)}, x_{11}^{(1)}; \dots; x_{00}^{(n)}, x_{01}^{(n)}, x_{10}^{(n)}, x_{11}^{(n)}).$$

Niestety przedstawiona w poprzednich podrozdziałach strategia dowodzenia istnienia pokrycia unimodularnego nie da się zastosować w ogólnym przypadku, gdy grafy mogą mieć cykle. Na przykład grafów cyklicznych nie da się uzyskać z dwóch grafów bez cykli sklejając wzdłuż pojedynczej krawędzi. Dodatkowo sklejanie wzdłuż dwóch krawędzi (lub większej ich liczby) rodzi nowe problemy. Jak pokazuje bowiem przykład do definicji A.3, w ogólnym przypadku produkt włóknisty dwóch wielościanów kratowych nad wielościanem, który nie jest simpleksem, nie musi być wielościanem kratowym.

Z punktu widzenia dalszych rozważań użyteczna może być następująca definicja.

Definicja 3.12 *Niech $P_1 = \text{conv}\{e_1, \dots, e_n\} \subset (M_1)_{\mathbb{R}}$, $P_2 = \text{conv}\{f_1, \dots, f_m\} \subset (M_2)_{\mathbb{R}}$ i $Q \subset \mathbb{Z}^k$ będą wielościanami kratowymi oraz $\ell_i : M_i \rightarrow \mathbb{Z}^k$ dwoma homomorfizmami krat takimi, że $\ell_i(P_i) = Q$. Wówczas kratowym produktem włóknistym nad Q będziemy nazywać $P_1 \boxtimes_{\ell_2} P_2 = \text{conv}\{e_i + f_j \mid \ell_1(e_i) = \ell_2(f_j)\} \subset P_1 \times P_2$.*

Pojęcie kratowego produktu włóknistego pozwala uzyskać ciekawą charakteryzację wielościanu dla modelu cyklicznego, która pomoże nam później lepiej go badać.

Konstrukcja 3.13 *Niech $R_{n+1} \subset (M_1)_{\mathbb{R}}$ i $R_3 \subset (M_2)_{\mathbb{R}}$ będą dwoma wielościanami kratowymi związanymi odpowiednio z modelem \mathcal{L}_{n+1} oraz \mathcal{L}_3 dla $n \geq 3$. Rozważmy dwa homomorfizmy krat $\ell_i : M_i \rightarrow \mathbb{Z}^6$ takie, że:*

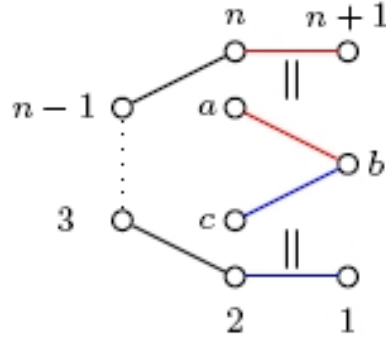
- $(\ell_1)_{\mathbb{R}}(R_{n+1}) = \Delta_3 \times \Delta_3$, gdzie Δ_3 to standardowy simpleks 3-wymiarowy

$$\ell_1(x) = [x_{01}^{(1)}, x_{10}^{(1)}, x_{11}^{(1)}; x_{01}^{(n)}, x_{10}^{(n)}, x_{11}^{(n)}],$$

- $(\ell_2)_{\mathbb{R}}(R_3) = \Delta_3 \times \Delta_3$

$$\ell_2(y) = [x_{01}^{(b)}, x_{10}^{(b)}, x_{11}^{(b)}; x_{01}^{(a)}, x_{10}^{(a)}, x_{11}^{(a)}].$$

Na poziomie grafów właściwy kratowy produkt włóknisty odpowiada poniższemu sklejeniu:



Wówczas kratowy produkt włóknisty $R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$ jest izomorficzny (w sensie homomorfizmów krat) z wielością $P_n \subset M_{\mathbb{R}}$ binarnego modelu cyklicznego.

Dowód: Wielościan $R_{n+1} \times R_3 \subset V_n \times V_2$ ma wierzchołki postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \cdots \oplus e_{i_n i_{n+1}}^{(n)} \oplus e_{j_a j_b}^{(a)} \oplus e_{j_b j_c}^{(b)},$$

gdzie $e_{ij}^{(k)}$ dla $k = 1, \dots, n$ to wektory jednostkowe przestrzeni V_n , a $e_{ij}^{(k)}$ dla $k = a, b$ to wektory jednostkowe przestrzeni V_2 oraz $(i_1, \dots, i_n) \in \{0, 1\}^{n+1}$, $(j_a, j_b, j_c) \in \{0, 1\}^3$.

Wielościan $R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$ z samej definicji kratowego produktu włóknistego ma wierzchołki postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \cdots \oplus e_{i_n i_1}^{(n)} \oplus e_{i_n i_1}^{(a)} \oplus e_{i_1 i_2}^{(b)}. \quad (3.12)$$

Rzeczywiście, $R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$ jest rozpinany przez te wierzchołki w produkcie wierzchołków R_{n+1} i R_3 , dla których zachodzi zgodność na homomorfizmach ℓ_1 i ℓ_2 . Równości

$$x_{01}^{(1)} = x_{01}^{(b)}, x_{10}^{(1)} = x_{10}^{(b)}, x_{11}^{(1)} = x_{11}^{(b)}; x_{01}^{(n)} = x_{01}^{(a)}, x_{10}^{(n)} = x_{10}^{(a)}, x_{11}^{(n)} = x_{11}^{(a)};$$

implikują równości $x_{00}^{(1)} = x_{00}^{(b)}$ oraz $x_{00}^{(n)} = x_{00}^{(a)}$. Wierzchołki w równaniu 3.12 to jedyne spośród wierzchołków $R_{n+1} \times R_3$, które wszystkie te osiem równości spełniają jednocześnie.

Z drugiej strony na mocy równości 3.11 łatwo zauważyć, że wielościany $R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$ oraz P_n są izomorficzne (jako wielościany kratowe). Przekształcenie $f : M_1 \ell_1 \boxtimes_{\ell_2} M_2 \rightarrow M$ zadane jest jako

$$f(e_{ij}^{(k)}) = e_{ij}^{(k)} \quad \text{dla } k \neq a, b$$

$$f(e_{ij}^{(a)}) = f(e_{ij}^{(b)}) = 0.$$

Przekształcenie w drugą stronę zadane jest jako

$$g(e_{ij}^{(k)}) = e_{ij}^{(k)} \quad \text{dla } k \neq 1, n$$

$$g(e_{ij}^{(1)}) = e_{ij}^{(1)} + e_{ij}^{(b)}; \quad g(e_{ij}^{(n)}) = e_{ij}^{(n)} + e_{ij}^{(a)}$$

a zatem istotnie z punktu widzenia równoważności krat, wystarczy badać wielościan $R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$. \square

3.3. Więcej o wielościanie P_n

W tym podrozdziale zajmiemy się wielościanem P_n w sposób bardziej kompleksowy. W poprzednim podrozdziale stwierdziliśmy, że wielościan ten jest normalny dla $n = 3, 4, 5, 6, 7$, a zatem w tych przypadkach definiuje semitoryczną rozmaitość rzutową, która jest normalna. W związku z tym rozmaitość X_{P_n} (patrz podrozdział 1.3) i Y_n , która jest zadana przez model \mathcal{C}_n , są sobie równe dla małych n . Jak pokazują rozważania przedstawione w rozdziale 1, wielościan ten zawiera zatem ważną informację o rozmaitości Y_n . Jest to główną motywacją dla dokładniejszej jego analizy. Innym ważnym powodem analizy tego wielościanu dla ogólnego n jest próba zrozumienia jego budowy, co może w przyszłości pomóc rozstrzygnąć czy wielościan ten jest normalny czy nie.

Wierzchołki wielościanu P_n będziemy często utożsamiać z ich współrzędnymi lub z ciągami zero-jedynkowymi $(i_1, \dots, i_n) \in \{0, 1\}^n$, którym odpowiadają. Będziemy zatem zależnie od kontekstu korzystać z jednego z dwóch przedstawień:

$$\{0, 1\}^n \ni (i_1, \dots, i_n) \sim e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \dots \oplus e_{i_{n-1} i_n}^{(n-1)} \oplus e_{i_n i_1}^{(n)} \in V_n$$

Wymiar wielościanu P_n przy wykorzystaniu argumentu kombinatorycznego i algebraicznego wyliczyli Hoşten i Sullivant [HS02]. Pokazali oni, że wymiar ten wynosi $2n$. Na mocy rozważań z podrozdziału 1.1, z powyższego faktu wynika w szczególności, że $\dim(Y_n) = 2n$.

Opis wielościanu P_n będzie przebiegał w kilku krokach. Będziemy korzystać z reprezentacji P_n jako kratowego produktu włóknistego, jak w konstrukcji 3.13. Rozumowanie podobne, jak w lemacie 3.6, pozwoli nam w prosty sposób znaleźć reprezentację za pomocą równości i nierówności dla wielościanu $R_{n+1} \ell_1 \times \ell_2 R_3$. Jednak, jak się okaże, wielościan ten nie jest wielościanem kratowym.

Przykład 3.14 *Wskażemy punkt wielościanu $R_4 \ell_1 \times \ell_2 R_3$, który nie należy do wielościanu kratowego $R_4 \ell_1 \boxtimes \ell_2 R_3$. Niech dane będą dwa wierzchołki produktu $R_4 \times R_3$ postaci:*

$$p_1 = (e_{00}^{(1)} \oplus e_{00}^{(2)} \oplus e_{01}^{(3)}) \oplus (e_{01}^{(a)} \oplus e_{11}^{(b)}),$$

$$p_2 = (e_{11}^{(1)} \oplus e_{11}^{(2)} \oplus e_{10}^{(3)}) \oplus (e_{10}^{(a)} \oplus e_{00}^{(b)}).$$

Niech $p = (\frac{1}{2}p_1 + \frac{1}{2}p_2)$. Wówczas

$$p = [\frac{1}{2}, 0, 0, \frac{1}{2}; \frac{1}{2}, 0, 0, \frac{1}{2}; 0, \frac{1}{2}, \frac{1}{2}, 0; 0, \frac{1}{2}, \frac{1}{2}, 0; \frac{1}{2}, \frac{1}{2}, 0; \frac{1}{2}, 0, 0, \frac{1}{2}],$$

a zatem jak łatwo sprawdzić jest zgodność na odpowiednich współrzędnych (patrz konstrukcja 3.13), więc $p \in \ker(\ell_1 - \ell_2)$. Korzystając z homomorfizmu krat f z dowodu konstrukcji 3.13 dostajemy punkt $p' \in V_3$ o współrzędnych $p' = [\frac{1}{2}, 0, 0, \frac{1}{2}; \frac{1}{2}, 0, 0, \frac{1}{2}; 0, \frac{1}{2}, \frac{1}{2}, 0]$. Rozważmy nierówność w V_3 postaci

$$x_{00}^{(1)} - x_{01}^{(2)} - x_{00}^{(3)} \leq 0.$$

Można sprawdzić, że każdy wierzchołek P_3 musi spełniać tę nierówność. Rzeczywiście, żeby wierzchołek nie spełniał jej musiałoby w szczególności być spełnione $x_{00}^{(1)} = 1$, ale łatwo sprawdzić, że wierzchołki $(0, 0, 0)$, $(0, 0, 1)$ spełniają ją (jako równość). W konsekwencji każdy punkt wielościanu P_3 musi spełniać tę nierówność. Jednak p' jej nie spełnia, a zatem $p' \notin P_3$. Korzystając z homomorfizmu g z dowodu konstrukcji 3.13 dostajemy, że $p \notin R_4 \ell_1 \boxtimes \ell_2 R_3$.

Powyższy przykład uogólnia się. W konsekwencji nie może zachodzić równość pomiędzy wielościanami $R_{n+1} \ell_1 \times \ell_2 R_3$ oraz $R_{n+1} \ell_1 \boxtimes \ell_2 R_3$. Mówi o tym poniższa uwaga.

Uwaga 3.15 Ustalmy ciąg $(j_1, j_2, \dots, j_n) \in \{0, 1\}^n$ i niech dane będą dwa wierzchołki produktu $R_{n+1} \times R_3$ postaci:

$$p_1 = (e_{j_1 j_2}^{(1)} \oplus e_{j_2 j_3}^{(2)} \oplus \dots \oplus e_{j_n 1-j_1}^{(n)}) \oplus (e_{j_n 1-j_1}^{(a)} \oplus e_{1-j_1 1-j_2}^{(b)}),$$

$$p_2 = (e_{1-j_1 1-j_2}^{(1)} \oplus e_{1-j_2 1-j_3}^{(2)} \oplus \dots \oplus e_{1-j_n j_1}^{(n)}) \oplus (e_{1-j_n j_1}^{(a)} \oplus e_{j_1 j_2}^{(b)}).$$

Przy oznaczeniach z poprzedniego podrozdziału niech p będzie kombinacją liniową powyższych wierzchołków taką, że $p = (\frac{1}{2}p_1 + \frac{1}{2}p_2) \in \ker(\ell_1 - \ell_2)$. Zachodzi zatem $p \in R_{n+1} \ell_1 \times_{\ell_2} R_3$, ale $p \notin R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$.

Dowód: To, że $p \in \ker(\ell_1 - \ell_2)$ jest oczywiste. Istotnie, łatwo spostrzec, że współrzędne bloku (n) są równe współrzędnym bloku (a) , a współrzędne bloku (1) , współrzędnym bloku (b) . Dokładniej

$$p = \frac{1}{2} \cdot \left((e_{j_1 j_2}^{(1)} + e_{1-j_1 1-j_2}^{(1)}) \oplus (e_{j_2 j_3}^{(2)} + e_{1-j_2 1-j_3}^{(2)}) \oplus \dots \oplus (e_{j_n 1-j_1}^{(n)} + e_{1-j_n j_1}^{(n)}) \oplus \right. \\ \left. \oplus (e_{j_n 1-j_1}^{(a)} + e_{1-j_n j_1}^{(a)}) \oplus (e_{1-j_1 1-j_2}^{(b)} + e_{j_1 j_2}^{(b)}) \right)$$

co po rzutowaniu na V_n , które przeprowadzało $R_{n+1} \ell_1 \boxtimes_{\ell_2} R_3$ na P_n (dowód konstrukcji 3.13) daje punkt p' równy

$$\frac{1}{2} \cdot \left((e_{j_1 j_2}^{(1)} + e_{1-j_1 1-j_2}^{(1)}) \oplus (e_{j_2 j_3}^{(2)} + e_{1-j_2 1-j_3}^{(2)}) \oplus \dots \oplus (e_{j_{n-1} j_n}^{(n-1)} + e_{1-j_{n-1} 1-j_n}^{(n-1)}) \oplus (e_{j_n 1-j_1}^{(n)} + e_{1-j_n j_1}^{(n)}) \right).$$

Ten punkt nie należy do wielościanu P_n . Aby to zobaczyć wystarczy rozważyć nierówność

$$x_{j_1 j_2}^{(1)} - x_{j_2 1-j_3}^{(2)} - \dots - x_{j_{n-1} 1-j_n}^{(n-1)} - x_{j_n j_1}^{(n)} \leq 0. \quad (3.13)$$

Zauważmy, że wszystkie wierzchołki P_n spełniają 3.13. Jedyna szansa by wierzchołek nie spełniał tej nierówności jest by $x_{j_1 j_2}^{(1)} = 1$ oraz reszta współrzędnych występujących w 3.13 się zerowała. Pierwszy warunek spełniają wierzchołki typu $(j_1 j_2 * \dots *)$. Kolejne warunki $x_{j_2 1-j_3}^{(1)} = \dots = x_{j_{n-1} 1-j_n}^{(n-1)} = 0$ jednoznacznie identyfikują spełniający je wierzchołek jako $(j_1 j_2 \dots j_n)$. Tylko, że ten wierzchołek spełnia $x_{j_n j_1}^{(n)} = 1$, a zatem dla niego

$$x_{j_1 j_2}^{(1)} - x_{j_2 1-j_3}^{(1)} - \dots - x_{j_{n-1} 1-j_n}^{(n-1)} - x_{j_n j_1}^{(n)} = 0.$$

W ten sposób pokazaliśmy, że wszystkie wierzchołki P_n spełniają 3.13, a zatem również wszystkie punkty P_n muszą tą nierówność spełniać. Łatwo jednak sprawdzić, że dla punktu p' nierówność ta nie jest spełniona, gdyż dla współrzędnych punktu p' zachodzi

$$x_{j_1 j_2}^{(1)} - x_{j_2 1-j_3}^{(1)} - \dots - x_{j_{n-1} 1-j_n}^{(n-1)} - x_{j_n j_1}^{(n)} = \frac{1}{2},$$

a zatem $p' \notin P_n$. W konsekwencji, przy zastosowaniu homomorfizmu g z dowodu konstrukcji 3.13 dostajemy tezę. \square

Kratowy produkt włóknisty dwóch wielościanów kratowych stanowi podwielościan produktu włóknistego rozpinany przez swoje punkty kratowe. Przejście do wielościanu $R_{n+1} \ell \boxtimes_{\ell_2} R_3$ polega na dodaniu do reprezentacji wielościanu $R_{n+1} \ell_1 \times_{\ell_2} R_3$ takich nierówności, które ograniczą go do kombinacji wypukłej swoich punktów kratowych. Jednym z wniosków z dowodu poprzedniej uwagi jest, że nierówność 3.13 wyznacza ścianę wielościanu P_n . Poniższy lemat mówi, że jest to ściana maksymalnego wymiaru.

Lemat 3.16 Niech $(j_1, j_2, \dots, j_n) \in \{0, 1\}^n$ oraz niech dana będzie nierówność

$$x_{j_1 j_2}^{(1)} - x_{j_2 1 - j_3}^{(2)} - \dots - x_{j_{n-1} 1 - j_n}^{(n-1)} - x_{j_n j_1}^{(n)} \leq 0. \quad (3.14)$$

Wówczas wyznacza ona simplicjalną ścianę maksymalnego wymiaru wielościanu P_n .

Dowód: Dowód poprzedniej uwagi pokazał, że wszystkie wierzchołki P_n spełniają tę nierówność, a wierzchołek (j_1, j_2, \dots, j_n) spełnia ją jako równość. Łatwo zauważyć, że wierzchołki

$$\begin{aligned} & (1 - j_1, j_2, j_3, \dots, j_n), \\ & (1 - j_1, 1 - j_2, j_3, \dots, j_n), \\ & \dots \\ & (1 - j_1, 1 - j_2, 1 - j_3, \dots, 1 - j_n), \\ & (j_1, 1 - j_2, 1 - j_3, \dots, 1 - j_n), \\ & \dots \\ & (j_1, j_2, \dots, j_{n-2}, 1 - j_n), \end{aligned}$$

również spełniają ją jako równość oraz, że dla każdego innego wierzchołka nierówność 3.14 jest ostra. Wymieniliśmy zatem $2n$ wierzchołków na ścianie wyznaczonej przez nierówność 3.14. Zauważmy dodatkowo, że spośród tych $2n$ wierzchołków dla każdego $k = 1, \dots, n$ dokładnie jeden wierzchołek spełnia $x_{j_k 1 - j_{k+1}} = 1$ oraz dokładnie jeden spełnia $x_{1 - j_k j_{k+1}} = 1$. Wobec tego wierzchołki te tworzą zestaw liniowo niezależnych wektorów. Skoro $\dim P_n = 2n$, to wierzchołki te wyznaczają simplicjalną ścianę maksymalnego wymiaru. Wszystkie ściany tego typu będziemy oznaczać przez (F2). \square

Przykład 3.17 Jeśli $(j_1, \dots, j_n) = (0, \dots, 0)$ mamy nierówność

$$x_{00}^{(1)} - x_{01}^{(2)} - \dots - x_{01}^{(n-1)} - x_{00}^{(n)} \leq 0,$$

którą spełnia wierzchołek $(0 \dots 0)$ oraz wszystkie wierzchołki (i_1, \dots, i_n) takie, że $i_1 = \dots = i_k = 0$ i $i_{k+1} = \dots = i_n = 1$ lub $i_1 = \dots = i_k = 1$ i $i_{k+1} = \dots = i_n = 0$ dla pewnego $k = 2, \dots, n$, gdzie $k = n$ odpowiada wierzchołkom $(0 \dots 0)$ i $(1 \dots 1)$.

Dowód lematu 3.16 pokazuje, że mamy 2^{n-1} różnych ścian simplicjalnych typu (F2). Kolejnym prostym wnioskiem, że nierówność

$$x_{i_1 i_2}^{(1)} - x_{i_2 1 - i_3}^{(2)} - \dots - x_{i_{n-1} 1 - i_n}^{(n-1)} - x_{i_n i_1}^{(n)} \leq 0$$

wyznacza tą samą ścianę, co nierówność

$$x_{1 - i_1 1 - i_2}^{(1)} - x_{1 - i_2 i_3}^{(2)} - \dots - x_{1 - i_{n-1} i_n}^{(n-1)} - x_{1 - i_n 1 - i_1}^{(n)} \leq 0$$

uzyskanej z poprzedniej poprzez zamianę wszystkich indeksów w subskryptach z zer na jedynki i na odwrót.

Z punktu widzenia informacji geometrycznej, którą niesie ze sobą wielościan P_n , istotna będzie jego konstrukcja w podprzestrzeni afinicznej, w której jest on już pełnowymiarowy. Użyteczna w tym kontekście będzie następująca uwaga.

Uwaga 3.18 *Wielościan P_n leży na $2n$ -wymiarowej podprzestrzeni afinicznej przestrzeni V_n opisywanej równościami*

$$\begin{aligned} x_{00}^{(i)} + x_{10}^{(i)} &= x_{00}^{(i+1)} + x_{01}^{(i+1)} \quad \text{dla } i = 1, \dots, n \\ x_{00}^{(i)} + x_{10}^{(i)} + x_{00}^{(i+1)} + x_{01}^{(i+1)} &= 1 \quad \text{dla } i = 1, \dots, n \end{aligned}$$

gdzie każdą, z pierwszych n równości można zamienić na równoważną

$$x_{01}^{(i)} + x_{11}^{(i)} = x_{10}^{(i+1)} + x_{11}^{(i+1)}.$$

Dowód powyższej uwagi polega na bezpośrednim sprawdzeniu, że wszystkie wierzchołki muszą spełniać każdą z wypisanych równości. Skoro $\dim P_n = 2n$, to są to już wszystkie równości opisujące podprzestrzeń afiniczną dla której wielościan P_n jest pełnego wymiaru.

Poniższa hipoteza dotyczy pełnego opisu wielościanu P_n dla dowolnego n w reprezentacji hiperpłaszczyznami podpierającymi. Przypuszczenie to pojawia się w postaci twierdzenia w artykule [HS02], jednak okazuje się, że ma ono niepoprawny dowód. Wynik ten jednak sformułujemy dlatego, że chcemy się do niego odwoływać później dla przypadku małych n , dla których hipoteza ta jest prawdziwa, o czym przekonują bezpośrednie obliczenia wykonane na komputerze. Warto dodać, że w cytowanej pracy ściany typu (F2) zostały uzyskane w inny sposób niż u nas i bez podawania ich jawnej postaci.

Hipoteza 3.19 *Wielościan P_n binarnego modelu cyklicznego C_n równy jest przecięciu podprzestrzeni afinicznych danych przez równania*

$$\begin{aligned} x_{00}^{(i)} + x_{10}^{(i)} &= x_{00}^{(i+1)} + x_{01}^{(i+1)} \\ x_{01}^{(i)} + x_{11}^{(i)} &= x_{10}^{(i+1)} + x_{11}^{(i+1)} \end{aligned} \quad \text{dla } 1 \leq i \leq n, \text{ oraz}$$

$$\sum_{i=1}^n (x_{00}^{(i)} + x_{10}^{(i)} + x_{01}^{(i)} + x_{11}^{(i)}) = n$$

z $4n$ ścianami maksymalnego wymiaru typu (F1) postaci $x_{j_i j_{i+1}}^{(i)} \geq 0$ dla $1 \leq i \leq n$ oraz 2^{n-1} ścianami typu (F2).

Uwaga 3.20 *Zestaw równości w hipotezie 3.19 można zamienić na równoważny:*

$$x_{00}^{(i)} + x_{10}^{(i)} = x_{00}^{(i+1)} + x_{01}^{(i+1)}$$

oraz

$$x_{00}^{(i)} + x_{10}^{(i)} + x_{01}^{(i)} + x_{11}^{(i)} = 1 \quad \text{dla każdego } i = 1, \dots, n.$$

Dowód: Jest to proste sprawdzenie, że z jednego zestawu równości wynika drugi i odwrotnie. \square

Wielościan z powyższej hipotezy będziemy dla odróżnienia oznaczać Q_n . Bezpośrednie obliczenia w programie `Polymake` pokazały, że dla $n = 3, 4, 5, 6, 7$ istotnie $P_n = Q_n$ (dla $n > 7$ obliczenia są na razie niemożliwe).

Dla pełniejszego wglądu w kombinatorykę wielościanu Q_n zauważmy jeszcze, że na każdej ścianie typu (F1) leży dokładnie $3 \cdot 2^{n-2}$ wierzchołków, a zatem ściana ta jest simplicjalna tylko dla przypadku, gdy $n = 3$.

Uwaga 3.18 (lub hipoteza 3.19) motywuje kolejne eliminacje zmiennych, tak aby rozważać wielościan dokładnie w przestrzeni afinicznej zdefiniowanej przez odpowiednie równania.

Operacja ograniczenia się do przestrzeni afinicznej $W_n = \text{span}\{e_{10}^{(i)}, e_{11}^{(i)}\}_{i=1}^n$ wymiaru $2n$ jest ważna, gdyż w tej sytuacji wektor normalny do ściany maksymalnego wymiaru staje się jednoznacznie wyznaczony z dokładnością do przemnożenia przez skalar. Będziemy z tego korzystać w naszych obliczeniach.

Zdefiniujemy przekształcenie

$$f : V_n \rightarrow W_n, \quad f(e_{00}^{(i)}) = f(e_{01}^{(i)}) = 0, f(e_{10}^{(i)}) = e_{10}^{(i)}, f(e_{11}^{(i)}) = e_{11}^{(i)}$$

dla wszystkich $i = 1, \dots, n$.

Jeżeli dla każdego i położymy $x_{00}^{(i)} = 1 - x_{10}^{(i)} - x_{01}^{(i)} - x_{11}^{(i)}$ to nie zmniejszymy liczby równości, które definiują P_n . Pozostaje n równości typu:

$$x_{01}^{(i)} + x_{11}^{(i)} = x_{10}^{(i+1)} + x_{11}^{(i+1)}.$$

Możemy kontynuować naszą redukcję kładąc

$$x_{01}^{(i)} = -x_{11}^{(i)} + x_{10}^{(i+1)} + x_{11}^{(i+1)}. \quad (3.15)$$

Pierwsze podstawienie zmienia się w ten sposób na:

$$x_{00}^{(i)} = 1 - x_{10}^{(i)} - x_{10}^{(i+1)} - x_{11}^{(i+1)}, \quad (3.16)$$

Dzięki tym dwóm ostatnim równaniom możemy łatwo zadać użyteczną parametryzację podprzestrzeni afinicznej, w której leży P_n przez przestrzeń W_n . Parametryzacja ta wraz z przekształceniem f zadaje izomorfizm pomiędzy P_n a jego włożeniem w P_n . Jest ona zadana wzorami

$$\begin{aligned} x_{00}^{(i)} &= 1 - x_{10}^{(i)} - x_{10}^{(i+1)} - x_{11}^{(i+1)}, \\ x_{01}^{(i)} &= -x_{11}^{(i)} + x_{10}^{(i+1)} + x_{11}^{(i+1)}, \\ x_{10}^{(i)} &= x_{10}^{(i)}, x_{11}^{(i)} = x_{11}^{(i)} \end{aligned} \quad (3.17)$$

dla $i = 1, \dots, n$. Przekształcenie f będzie wykorzystywane w ostatniej części tej pracy, w której będziemy przeprowadzać dokładne obliczenia dla wielościanów P_n oraz wachlarzy normalnych Σ_n dla małych n .

Włożenie wielościanu P_n w podprzestrzeń W_n generowaną przez $\{e_{10}^{(i)}, e_{11}^{(i)}\}_{i=1}^n$, zadane przekształceniem f , stanowi wygodną operacją, musimy jednak pamiętać, że wielościan P_n zawiera w sobie pewną geometryczną informację. Potrzebujemy pewności, że jej nie stracimy (patrz definicja 1.13). W naszym przypadku konstrukcja przekształcenia odwrotnego g pokazuje, że w istocie P_n jest kratowo równoważny ze swoją kopią włożoną w podprzestrzeń W_n .

Lemat 3.21 *Przekształcenie $f : V_n \rightarrow W_n$ zdefiniowane przed chwilą zadaje izomorfizm między kratą generowaną przez wielościan P_n , a kratą generowaną przez zbiór wektorów $\{e_{10}^{(i)}, e_{11}^{(i)}\}_{i=1}^n$.*

Dowód: Skoro $f(P_n)$ jest 0/1-wielościanem, to na mocy wcześniejszych rozważań, za pomocą pewnych kombinacji całkowitoliczbowych elementów wymienionych w lemacie potrafimy w sposób jednoznaczny wyznaczyć każdy z wierzchołków tego wielościanu. Jednoznaczność wynika z tego, że wektory te stanowią bazę wektorową dla W_n . Ale f wyznacza izomorfizm krat. \square

Teraz opiszemy strukturę wielościanu P_n , jako pełnowymiarowego wielościanu w przestrzeni W_n . Zauważmy, że inaczej niż przypadku $4n$ -wymiarowym w nierównościach wyznaczających ściany maksymalnego wymiaru może pojawiać się niezerowy współczynnik, który będziemy oznaczać przez a_F . Wynika to z postaci przekształcenia $g^* : V_n^* \rightarrow W_n^*$, które zdefiniowaliśmy równaniem 3.17.

Lemat 3.22 *Dla modelu C_n w dokładnie n różnych nierównościach typu (F2) po włożeniu w przestrzeń W_n mamy $a_F = 0$.*

Dowód: Nierówność typu

$$x_{i_1 i_2}^{(1)} - x_{i_2 1 - i_3}^{(2)} - \dots - x_{i_{n-1} 1 - i_n}^{(n-1)} - x_{i_n i_1}^{(n)} \leq 0$$

przekształcamy na nierówność na przestrzeni W_n za pomocą g^* z równania 3.17. Widać, że niezerowe współczynniki związane są z $x_{00}^{(i)} \in V_n^*$. Z postaci tej nierówności wynika, że zerowe współczynniki pojawią się tylko dla takich $(i_1, \dots, i_n) \in \{0, 1\}^n$, że

- albo $(i_1, i_2) \neq (0, 0)$, $(i_n, i_1) \neq (0, 0)$ oraz $(i_k, i_{k+1}) \neq (0, 1)$ dla wszystkich $k = 2, \dots, n-1$
- albo $(i_1, i_2) = (0, 0)$, $(i_n, i_1) \neq (0, 0)$ oraz $(i_k, i_{k+1}) = (0, 1)$ dla dokładnie jednego $k = 2, \dots, n-1$
- albo $(i_1, i_2) = (0, 0)$, $(i_n, i_1) = (0, 0)$ oraz $(i_k, i_{k+1}) \neq (0, 1)$ dla wszystkich $k = 2, \dots, n-1$.

Pierwszy warunek spełnia n ciągów zero-jedynkowych postaci $i_1 = \dots = i_j = 1$ oraz $i_{j+1} = \dots = i_n = 0$ dla $j = 1, \dots, n$ (to znaczy, że ciąg jedynek też wliczamy). Drugi warunek spełniają ciągi postaci $i_1 = \dots = i_j = 0$ oraz $i_{j+1} = \dots = i_n = 1$ dla $j = 2, \dots, n-1$. Trzeci warunek, jak łatwo sprawdzić, spełnia jedynie ciąg zerowy. Pamiętając jednak, że ciąg (i_1, i_2, \dots, i_n) zadaje tą samą nierówność co ciąg $(1 - i_1, 1 - i_2, \dots, 1 - i_n)$ (patrz poprzedni podrozdział) dostajemy tezę. \square

Wniosek 3.23 *Spośród $2^{n-1} + 4n$ nierówności wymienionych w hipotezie 3.19 dokładnie n nierówności typu (F1) oraz $2^{n-1} - n$ nierówności typu (F2) posiada niezerowy wyraz wolny.*

Dowód: Łatwo sprawdzić, że $4n$ nierówności typu (F1) jest przez g^* przekształcanych na:

$$-x_{10}^{(i)} - x_{10}^{(i+1)} - x_{11}^{(i+1)} \geq -1 \quad (3.18)$$

$$-x_{11}^{(i)} + x_{10}^{(i+1)} + x_{11}^{(i+1)} \geq 0 \quad (3.19)$$

$$x_{10}^{(i)} \geq 0 \quad (3.20)$$

$$x_{11}^{(i)} \geq 0. \quad (3.21)$$

Czyli dokładnie n z nich ma niezerowy współczynnik przy nierówności. Przypadek pozostałych ścian jest treścią lematu 3.22. \square

Dowód powyższych wyników pokazuje, że przy włożeniu odrobiny pracy można pokazać ile z nierówności typu (F2) ma współczynniki a_F wyższe niż jeden. Maksymalna wartość tego współczynnika wynosi $\lceil \frac{n}{2} \rceil - 1$ i odpowiada ciągowi $(0, 1, 0, 1, \dots, 0, 1)$ w przypadku, gdy n jest parzyste lub $(1, 0, 1, 0, \dots, 0, 1)$ w przypadku, gdy n jest nieparzyste.

3.4. Rozmaitość Y_n modelu \mathcal{C}_n dla małych n

W tym podrozdziale przedstawimy podstawowe własności rozmaitości modeli \mathcal{C}_n dla małej liczby wierzchołków grafu. Wpierw jednak przedstawimy kilka podstawowych wyników dla przypadku ogólnego. Będziemy analizować wielościan opisany nierównościami z hipotezy 3.19 oznaczając go przez Q_n . Treść hipotezy 3.19 mówi, że $Q_n = P_n$ i obliczenia pokazują, że jest to prawdziwe dla $n = 3, 4, 5, 6, 7$. Dlatego analiza Q_n pozwoli nam w sposób możliwie ogólny opisać wachlarz Σ_n dla małych n .

Wielościan Q_n ma $4n + 2^{n-1}$ ścian maksymalnego wymiaru. Aby uzyskać jednoznaczność w wyznaczaniu wektorów normalnych do tych ścian, skorzystaliśmy z odpowiedniego włożenia w podprzestrzeń $W_n = \text{span}\{e_{10}^{(i)}, e_{11}^{(i)}\}_{i=1}^n$. Wiemy również, że w sensie wielościanów kratowych, dostajemy taką samą informację ograniczając się do tego pełnowymiarowego przypadku. Zatem możemy zidentyfikować interesujące nas wektory normalne do ścian wielościanu Q_n jako $4n$ wektory pochodzące od ścian typu (F1) $x_{ij}^{(k)} \geq 0$ oraz 2^{n-1} wektory typu (F2), czyli:

- $2n$ wektory jednostkowe w przestrzeni W_n , które będziemy oznaczać przez $\epsilon_{ij}^{(k)}$.
- n wektorów typu $-\epsilon_{10}^{(k)} - \epsilon_{10}^{(k+1)} - \epsilon_{11}^{(k+1)}$ dla $k = 1, \dots, n$, gdzie w definiującej nierówności jest wyraz wolny równy 1.
- n wektorów typu $-\epsilon_{11}^{(k)} + \epsilon_{10}^{(k+1)} + \epsilon_{11}^{(k+1)}$ dla $k = 1, \dots, n$.
- oraz 2^{n-1} wektorów normalnych do ścian typu (F2),

gdzie ϵ oznacza wektor sprzężony do wektora e . Powyższe wektory w przestrzeni dualnej W_n^* generują 2^n stożków maksymalnego wymiaru (bo tyle jest wierzchołków w wielościanie). Wektory normalne do ścian typu (F2) jest stosunkowo ciężko opisać w sposób systematyczny. Wiemy na przykład, że nierówności

$$x_{11}^{(1)} - x_{10}^{(2)} - \dots - x_{10}^{(n-1)} - x_{11}^{(n)} \leq 0$$

odpowiada wektor normalny w W_n postaci

$$n_0 = [0, -1; 1, 0; 1, 0; \dots; 1, 0; 0, 1]$$

gdzie średnik oddziela każdy z n bloków tej przestrzeni.

Każdy wierzchołek wielościanu Q_n leży w $3n$ ścianach typu (F1), oraz w n ścianach typu (F2). Pierwsze jest jasne. Aby przekonać się co do tego drugiego stwierdzenia, z symetrii zadania, starczy pokazać, że $p_0 = (00 \dots 00)^1$ leży w n ścianach typu (F2), ale to można pokazać analizując lemat 3.23. Zatem każdy stożek maksymalnego wymiaru w wachlarzu Σ_n (patrz podrozdział 1.2) generowany jest przez $4n$ wektorów, z czego n pochodzi od ścian typu (F2).

W poniższych paragrafach zajmiemy się szczególnymi przypadkami rozmaitości Y_n dla $n = 3, 4$. W tym wypadku mamy $P_n = Q_n$. Wiedząc dodatkowo, że rozmaitość Y_n jest normalna dla każdego $n \geq 3$, wiemy również, że $Y_n = X_{P_n}$, gdzie konstrukcja rozmaitości X_P z wielościanu P została przedstawiona w podrozdziale 1.2. Analizując geometrię rozmaitości Y_n możemy zatem w szczególności korzystać z rozważań przedstawionych w podrozdziale 1.3.

¹Przypominamy o konwencji utożsamiania wierzchołków z ciągami zer i jedynek. W tym wypadku ciągowi $(00 \dots 00)$ odpowiada punkt $(1, 0, 0, 0; \dots; 1, 0, 0, 0)$ w V_n oraz punkt $(0, 0; \dots; 0, 0)$ w W_n .

Rozmaitość Y_3 Gdy $n = 3$, Y_3 ma wymiar sześć i jako, że zadana jest poprzez parametryzację w przestrzeni \mathbb{P}^7 , to musi być hiperpowierzchnią. Stosunkowo łatwo można sprawdzić, że równanie, które ją generuje to:

$$p_{000}p_{011}p_{101}p_{110} - p_{001}p_{010}p_{100}p_{111} = 0, \quad (3.22)$$

gdzie $\{p_{ijk}\}$ tworzą współrzędne jednorodne przestrzeni \mathbb{P}^7 . Przyjrzymy się teraz bliżej wachlarzowi definiującemu tą rozmaitość. Zbiór współrzędnych wierzchołków wielościanu P_n to

$$\{e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus e_{i_3 i_1}^{(3)} \mid (i_1, i_2, i_3) \in \{0, 1\}^3\} \subset V_3.$$

Korzystając w transformacji zaprezentowanej w podrozdziale 3.3, będziemy te punkty zapisywać w przestrzeni $W_3 = \mathbb{R}^6$.

Na mocy rozważań z podrozdziału 1.3 wierzchołki wielościanu P_3 odpowiadają stożkom maksymalnego wymiaru w wachlarzu Σ_3 oraz punktom stałym działania torusa. Wachlarz Σ_3 złożony jest zatem z ośmiu stożków maksymalnego wymiaru, z których każdy rozpinany jest przez 12 wektorów. W sumie mamy 16 wektorów definiujących wachlarz rozmaitości odpowiadającej modelowi \mathcal{C}_3 , które odpowiadają wektorom normalnym do ścian analizowanego wielościanu. Jest to sześć wektorów jednostkowych, sześć pozostałych wektorów od ścian typu (F1), oraz cztery wektory od ścian (F2): $n_0 = [0, -1, 1, 0, 0, 1]$ oraz

$$[1, 0, 0, 1, 0, -1][0], \quad [0, 1, 0, -1, 1, 0][0], \quad [-1, 0, -1, 0, -1, 0][1],$$

gdzie w nawiasach kwadratowych dorzuciliśmy współczynniki, które pojawiały się w nierównościach definiujących ściany maksymalnego wymiaru. Współczynniki były niezerowe dokładnie w tych nierównościach, których nie spełnia punkt $q_0 = (000)$.

Stożek $\sigma_{000} \subset N_{\mathbb{R}}$ generowany jest przez 12 wektorów normalnych przez nierówności, które spełnia wierzchołek $q_0 = (000)$, czyli 9 wektorów od ścian leżących na podprzestrzeniach przestrzeni V_n takich że $x_{01}^{(k)} = 0$, $x_{10}^{(k)} = 0$, $x_{11}^{(k)} = 0$ dla $k = 1, 2, 3$ oraz 3 wektorów typu (F2) (wszystkie poza tym z niezerowym współczynnikiem). Łatwo zauważyć, że wektory te leżą na tej samej podprzestrzeni afinicznej $\{x \in \mathbb{R}^6 \mid \langle m, x \rangle = 1\}$, gdzie $m = [1, 1, 1, 1, 1, 1] \in (W_3)^*$. Podobnie minimalne generatory kratowe dla wszystkich innych stożków wachlarza Σ_3 leżą na wspólnych podprzestrzeniach afinicznych typu $\langle m, x \rangle = 1$.

Wielościan P_3 posiada triangulację unimodularną. Można jej dokonać na dwa sposoby. Jednym z nich jest na przykład podział na cztery simpleksy, jak w dodatku B.2.1. Triangulację tę możemy znaleźć na mocy rozważań geometrycznych. Zauważmy, że każda ze ścian wielościanu P_3 jest simplicjalna. Wyróżnijmy wierzchołek $q_0 = (000) = [1, 0, 0, 0; 1, 0, 0, 0; 1, 0, 0, 0] \in V_3$. Wierzchołek ten nie leży dokładnie w czterech ścianach maksymalnego wymiaru. Jeżeli dodamy ten wierzchołek do każdego z sześćcioelementowych zbiorów wierzchołków definiujących te ściany, to dostaniemy cztery simpleksy unimodularne tworzące triangulację P_3 .

Rozmaitość Y_4 W przypadku $n = 4$ sytuacja już się trochę komplikuje. Wachlarz Σ_4 złożony jest z 16 stożków maksymalnego wymiaru. Każdy z tych stożków ma 16 generatorów. Wielościan P_4 posiada 24 ściany maksymalnego wymiaru, z czego jedynie osiem jest simplicjalnych.

Podobnie, jak w poprzednim przypadku sprawdzimy, czy minimalne generatory kratowe stożka σ_{0000} leżą na wspólnej podprzestrzeni afinicznej typu $\langle m, x \rangle = 1$. Zauważmy, że normalne do ścian $x_{10}^{(k)} = 0$ oraz $x_{10}^{(k)} = 0$ zapewniają, że jeśli takie $m \in (W_4)^*$ istnieje to musi być wektorem jedynek. Wówczas jednak $\langle m, [0, -1; 1, 0; 1, 0; 0, 1] \rangle = 2$, co przeczy, że

takie m istnieje. Przy pewnym nakładzie pracy będziemy mogli pokazać, że dla dowolnego stożka z wachlarza Σ_4 , jego minimalne generatory kratowe nie mogą leżeć na wspólnej hiperpłaszczyźnie. Zauważmy, że podobnie będzie w przypadku $n = 5$ i wyższych.

Wielościan P_4 posiada triangulację unimodularną złożoną z 64 simpleksów, które wypisane zostały w dodatku B.2.2. Równanie 3.6 parametryzuje toryczną rozmaitość rzutową w \mathbb{P}^{15} , która dana jest przez siedem równań²:

$$\begin{aligned}
p_{0000}p_{0101} - p_{0001}p_{0100}, \\
p_{0010}p_{1000} - p_{0000}p_{1010}, \\
p_{0010}p_{0111} - p_{0011}p_{0110}, \\
p_{1000}p_{1101} - p_{1001}p_{1100}, \\
p_{0110}p_{1100} - p_{0100}p_{1110}, \\
p_{0101}p_{1000}p_{1111} - p_{0111}p_{1001}p_{1100}, \\
p_{0000}p_{0101}p_{0110}p_{1011} - p_{0010}p_{0100}p_{0111}p_{1001},
\end{aligned}$$

gdzie p_{ijkl} tworzą współrzędne jednorodne na \mathbb{P}^{15} .

²Powyższe obliczenia zostały wykonane za pomocą programu `4ti2` [HHM05].

Rozdział 4

Zakończenie

Możemy wymienić dwa główne cele powyższej pracy. Z jednej strony chcieliśmy przedstawić podstawowe idee związane ze statystyką algebraiczną. Pokazaliśmy dlaczego w większości przypadków nie tracimy ogólności rozważań rozpatrując modele statystyczne jako obrazy przekształceń wielomianowych. Motywuje to wykorzystanie narzędzi geometrii algebraicznej w analizie statystycznej. Wedle naszej najlepszej wiedzy jest to pierwsze opracowanie w języku polskim poruszające ten temat. Wierzymy, że nie jest ono ostatnim takim. Rosnące moce obliczeniowe komputerów oraz pojawianie się nowych algorytmów algebry obliczeniowej (motywowanych mnogością ciekawych zastosowań) sprawia, że takie podejście do modelowania statystycznego ma przed sobą prawdopodobnie przyszłość.

Praca ta jednak idzie dalej wpisując się w nurt, który zapoczątkowali po części Bernd Sturmfels z Sethem Sullivantem, a po części Jarosław Wiśniewski z Weroniką Buczyńską. Podejście to opiera się na gruntownej analizie geometrii abstrakcyjnych rozmaitości, które pojawiają w rozważaniach statystyki algebraicznej. Ciężko na tym etapie badań powiedzieć zdecydowanie, czy tak abstrakcyjne podejście przyniesie jakieś korzyści w aplikacjach. Wierzymy jednak, że zagadnienia te są warte dalszych badań z dwóch powodów. Z jednej strony, co pokazała na przykład praca Buczyńskiej i Wiśniewskiego [BW06], analiza tych rozmaitości może prowadzić do ciekawych rodzin rozmaitości o interesujących właściwościach, a zatem do pogłębienia naszej wiedzy z dziedziny geometrii algebraicznej. Po drugie, z naszego punktu widzenia niezwykle ważne jest zbieranie wszelkich informacji o analizowanych modelach statystycznych. Wiedza ta może posłużyć nam w przyszłości do dogłębnego ich poznania.

Właśnie w tym kontekście należy wymienić główny wynik naszej pracy. Polega on na udowodnieniu, że rozmaitości związane z binarnymi modelami grafów bez cykli są normalne, a zatem są toryczne w klasycznym sensie. Jest to przybliżenie się do hipotezy podniesionej przez Sturmfelsa i Sullivanta w [SS06], że normalne są rozmaitości dla binarnych modeli grafów bez podgrafów postaci K_4 , tzn. grafów pełnych o czterech wierzchołkach. Następnym krokiem jest próba możliwie dokładnego opisu przypadku cyklicznego. W tej pracy postanowiliśmy oprzeć się na pracy Hoştena i Sullivanta [HS02] i rozszerzyć ją o nowe wyniki. Pozwoliło nam na to spojrzenie na wielościan modelu cyklicznego jako na kratowy podwielościan produktu włóknistego dwóch wielościanów dla modeli liniowych. Możemy zatem wyniki tej pracy potraktować jako dobry punkt wyjścia do próby dowiedzenia przypuszczenia Sturmfelsa i Sullivanta. Zauważmy ponadto, że wykazanie, że rozmaitości dla drzew są normalne otwiera również sposobność do dokładniejszej ich analizy klasycznymi metodami geometrii torycznej. Będziemy wówczas w stanie lepiej zrozumieć osobliwości tych rozmaitości oraz inne ważne geometryczne cechy.

Dodatek A

Lematy o unimodularności

W tym dodatku na początku przedstawimy za Buczyńską i Wiśniewskim [BW06] podstawowe wyniki dotyczące istnienia unimodularnego pokrycia dla wielościanu kratowego w pewnych szczególnych przypadkach. Następnie rozszerzymy je o serię lematów i twierdzenie, użyteczne w kontekście binarnych modeli grafów.

Pożyteczny będzie następujący lemat.

Lemat A.1 (O małym zaburzeniu) *Niech dane będą dwa pełnowymiarowe wielościany kratowe*

$$Q_1 = \text{conv}(e_1, \dots, e_r) \subset \mathbb{R}^m, \quad Q_2 = \text{conv}(f_1, \dots, f_s) \subset \mathbb{R}^n$$

oraz $(n+m)$ -wymiarowy zbiór domknięty w ich produkcie $F \subset Q_1 \times Q_2 \subset \mathbb{R}^{n+m}$. Wówczas jeśli $x = \sum_{i=1}^r a_i e_i + \sum_{j=1}^s b_j f_j \in \mathbb{R}^{n+m}$ nie należy do F to nie należy do F również dostatecznie małe zaburzenie punktu x . W związku z tym badając, czy $x \in F$ możemy założyć, że wszystkie a_i i b_j są różne od zera oraz że każde dwa podzbiory zbioru współczynników $\{a_i\}$ i $\{b_j\}$ mają różne sumy.

Dowód: F jest domkniętym podzbiorem $Q_1 \times Q_2$, a zatem $U = (Q_1 \times Q_2) \setminus F$ jest podzbiorem otwartym. Punkt $x \in Q_1 \times Q_2$ należy do U wtedy i tylko wtedy, gdy każde dostatecznie małe otoczenie punktu x zawiera się w U . Teza wynika z faktu, że $x \notin F$ wtedy i tylko wtedy, gdy $x \in U$. \square

Przedstawimy wpiery lemat, którego dowód można znaleźć w [BW06]. Prezentujemy go jednak dla zachowania przejrzystości rozumowania w dowodach następujących lematów.

Lemat A.2 *Niech $\Delta_1 \subset (M_1)_{\mathbb{R}}$ i $\Delta_2 \subset (M_2)_{\mathbb{R}}$ będą dwoma simpleksami unimodularnymi. Wówczas $\Delta_1 \times \Delta_2$ ma pokrycie unimodularne w kratce $M_1 \times M_2$.*

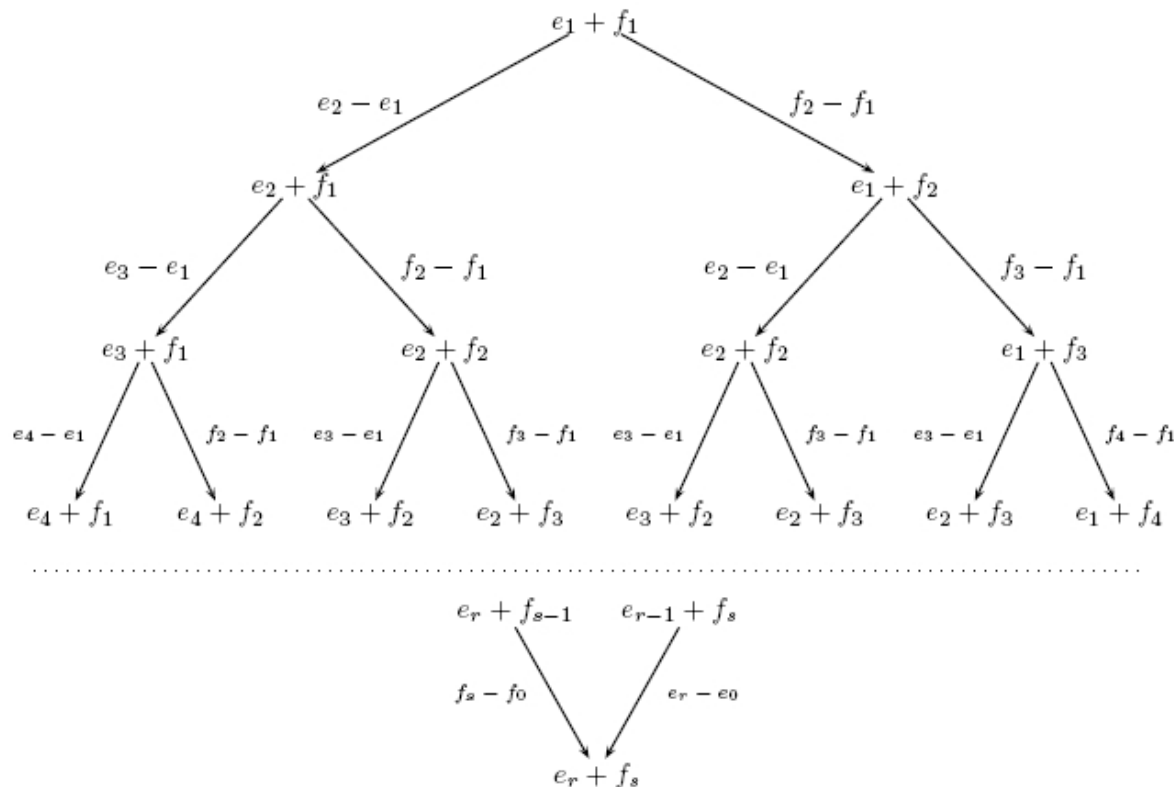
Dowód: Możemy założyć, że Δ_1 ma wierzchołki e_1, e_2, \dots, e_r , a Δ_2 ma wierzchołki f_1, f_2, \dots, f_s . Na mocy unimodularności tych simpleksów, krata M_1 generowana jest przez wektory $e_2 - e_1, \dots, e_r - e_1$, a krata M_2 przez wektory $f_2 - f_1, \dots, f_s - f_1$ (patrz podrozdział 1.2.3). Niech $x \in \Delta_1 \times \Delta_2$ będzie dane jako:

$$x = \sum_{i=1}^r a_i e_i + \sum_{j=1}^s b_j f_j,$$

gdzie $a_i, b_j \geq 0$ oraz $\sum a_i = \sum b_j = 1$.

Suma unimodularnych simpleksów zawartych w $\Delta_1 \times \Delta_2$ jest domkniętym podzbiorem w produkcie. Z lematu o małym zaburzeniu możemy założyć, że wszystkie współczynniki a_i i

Dodatkowo przy zastosowaniu prostych operacji można pokazać, że zbiór wektorów $\mathcal{B} = \{(e_i + f_j) - (e_1 + f_1) : c_{i,j} \neq 0, (i, j) \neq (1, 1)\}$ da się sprowadzić do standardowej bazy kraty $M_1 \times M_2$. Wystarczy na początek zauważyć, że w powyższym zbiorze wektorów znajduje się któryś z wektorów $(e_2 + f_1) - (e_1 + f_1) = e_2 - e_1$ lub $(e_1 + f_2) - (e_1 + f_1) = f_2 - f_1$.



Rekurencyjna procedura, której schemat pojawia się na powyższym diagramie (zacięzione kwadraty na rysunku A.1 mówią, jak poruszać się po tym diagramie) zapewni, że \mathcal{B} generuje wszystkie wektory $e_i - e_1$ dla $i = 2, \dots, r$ oraz $f_i - f_1$ dla $i = 2, \dots, s$, a zatem bazę kraty $M_1 \times M_2$. Ale skoro $|\mathcal{B}| = r + s - 2$ to \mathcal{B} jest również minimalnym zbiorem generatorów, co kończy dowód. \square

Powyższy lemat ma swoje uogólnienie na produkty włókniste nad simpleksami kratowymi.

Definicja A.3 Niech $Q_1 \subset (M_1)_{\mathbb{R}}$ i $Q_2 \subset (M_2)_{\mathbb{R}}$ będą dwoma wielościanami kratowymi. Rozważmy dwa homomorfizmy krat $\ell_i : M_i \rightarrow \mathbb{Z}^k$ takie, że $\ell_1(Q_1) = \ell_2(Q_2) = Q$. Wówczas produktem włóknistym Q_1 i Q_2 nad Q nazywamy wielościan $Q_1 \ell_1 \times_{\ell_2} Q_2$ taki, że

$$Q_1 \ell_1 \times_{\ell_2} Q_2 = (Q_1 \times Q_2) \cap \ker(\ell_1 - \ell_2).$$

Odpowiedni produkt włóknisty krat $M_1 \ell_1 \times_{\ell_2} M_2$ to krata

$$M_1 \ell_1 \times_{\ell_2} M_2 = (M_1 \times M_2) \cap \ker(\ell_1 - \ell_2).$$

W ogólności nie ma powodu by oczekiwać, że produkt włóknisty dwóch wielościanów kratowych jest wielościanem kratowym w kracie $M_1 \ell_1 \times_{\ell_2} M_2$. Zilustruje to poniższy przykład.

Przykład A.4 Niech $\Delta_1 = \Delta_2 = \text{conv}\{[0, 0, 0], [1, 1, 0], [1, 0, 1], [0, 1, 1]\}$ i rozważmy $\ell_i : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ będące rzutowaniem na dwie pierwsze współrzędne. Obliczenia pokazują, że

$$\Delta_1 \ell_1 \times_{\ell_2} \Delta_2 = \text{conv}\{[0, 0, 0, 0, 0, 0], [1, 1, 0, 1, 1, 0], [1, 0, 1, 1, 0, 1], [0, 1, 1, 0, 1, 1], [\frac{1}{2}, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{2}, 0], [\frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}]\}.$$

Krata $M_1 \times_{\ell_1} M_2$ to podkrata w \mathbb{Z}^6 taka, że $x_1 = y_1$ oraz $x_2 = y_2$, gdzie $x_1, x_2, x_3, y_1, y_2, y_3$ to współrzędne w \mathbb{R}^6 . Jak widać, produkt włóknisty w tym przypadku ma dodatkowo dwa wierzchołki spoza kraty.

Pojawia się naturalne pytanie o to, kiedy produkt włóknisty będzie miał wszystkie wierzchołki w kracie. Okazuje się, że mamy to zapewnione w jednym z interesujących przypadków, o których mówi poniższy lemat będący uogólnieniem lematu A.3 z [BW06]. Fakt, że w tym wypadku produkt włóknisty jest wielościanem kratowym wynika z samego dowodu i jest treścią uwagi A.7. Wpierw jednak przedstawimy modyfikację lematu o małym zaburzeniu, która będzie użyteczna w następnych dowodach.

Lemat A.5 (O małym zaburzeniu, drugi) *Niech dane będą dwa pełnowymiarowe wielościany kratowe*

$$Q_1 = \text{conv}(e_1, \dots, e_r) \subset \mathbb{R}^{m_1}, \quad Q_2 = \text{conv}(f_1, \dots, f_s) \subset \mathbb{R}^{m_2}$$

oraz homomorfizmy krat $\ell_i : \mathbb{Z}^{m_i} \rightarrow \mathbb{Z}^n$ takie, że $\ell_i(Q_i) = Q$ dla pewnego n -wymiarowego wielościanu kratowego. Niech $F \subset Q_1 \times_{\ell_1} Q_2 \subset \mathbb{R}^n$ będzie n -wymiarowym zbiorem domkniętym w produkcie włóknistym nad Q . Wówczas jeśli $x = \sum_{i=1}^r a_i e_i + \sum_{j=1}^s b_j f_j \in \mathbb{R}^n$ nie należy do F to nie należy do F również dostatecznie małe zaburzenie punktu x . W związku z tym badając, czy $x \in F$ możemy założyć, że wszystkie a_i i b_j są różne od zera oraz że każde dwa podzbiory zbioru współczynników $\{a_i\}$ i $\{b_j\}$ mają różne sumy. Jedynie równości, które są niezmiennicze to te, które zapewniają, że $x \in \ker(\ell_1 - \ell_2)$.

Dowód: Dowód przebiega analogicznie, jak dowód pierwszego lematu o małym zaburzeniu. Topologię na $\ker(\ell_1 - \ell_2)$ mamy indukowaną z topologii $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$. \square

Lemat A.6 *Niech $\Delta_1 \subset (M_1)_{\mathbb{R}}$ i $\Delta_2 \subset (M_2)_{\mathbb{R}}$ będą dwoma simpleksami unimodularnymi. Rozważmy dwa homomorfizmy krat $\ell_i : M_i \rightarrow \mathbb{Z}^k$ takie, że $\ell_i(\Delta_i) = \Delta$, gdzie Δ jest k -wymiarowym standardowym simpleksem kratowym. Wówczas produkt włóknisty $\Delta_1 \times_{\ell_1} \Delta_2$ ma pokrycie unimodularne względem produktu włóknistego krat $M = M_1 \times_{\ell_1} M_2$.*

Dowód: Dowód przeprowadzimy tylko dla $k = 2$, jednak łatwo zauważyć, że dowód ogólny polega tylko na staranniejszym wypisywaniu odpowiednich indeksów.

Możemy założyć, że Δ_1 ma wierzchołki

$$e_1^{00}, \dots, e_{r_{00}}^{00}, e_1^{01}, \dots, e_{r_{01}}^{01}, e_1^{10}, \dots, e_{r_{10}}^{10}$$

a Δ_2 odpowiednio

$$f_1^{00}, \dots, f_{s_{00}}^{00}, f_1^{01}, \dots, f_{s_{01}}^{01}, f_1^{10}, \dots, f_{s_{10}}^{10},$$

gdzie $\ell_1(e_i^{00}) = \ell_2(f_j^{00}) = (0, 0)$, $\ell_1(e_i^{01}) = \ell_2(f_j^{01}) = (0, 1)$ oraz $\ell_1(e_i^{10}) = \ell_2(f_j^{10}) = (1, 0)$ dla dowolnych i, j . Załóżmy, że $x \in \Delta_1 \times_{\ell_1} \Delta_2$ przedstawia się jak poniżej

$$x = \sum_{i=1}^{r_{00}} a_i^{00} e_i^{00} + \sum_{i=1}^{r_{01}} a_i^{01} e_i^{01} + \sum_{i=1}^{r_{10}} a_i^{10} e_i^{10} + \sum_{j=1}^{s_{00}} b_j^{00} f_j^{00} + \sum_{j=1}^{s_{01}} b_j^{01} f_j^{01} + \sum_{j=1}^{s_{10}} b_j^{10} f_j^{10},$$

gdzie $a_i^{00}, a_i^{01}, a_i^{10}, b_j^{00}, b_j^{01}, b_j^{10} \geq 0$, $\sum_{i=1}^{r_{00}} a_i^{00} + \sum_{i=1}^{r_{01}} a_i^{01} + \sum_{i=1}^{r_{10}} a_i^{10} = 1$, $\sum_{j=1}^{s_{00}} b_j^{00} + \sum_{j=1}^{s_{01}} b_j^{01} + \sum_{j=1}^{s_{10}} b_j^{10} = 1$ oraz dodatkowo $\sum_{i=1}^{r_{01}} a_i^{01} = \sum_{j=1}^{s_{01}} b_j^{01}$ i $\sum_{i=1}^{r_{10}} a_i^{10} = \sum_{j=1}^{s_{10}} b_j^{10}$, a w konsekwencji także $\sum_{i=1}^{r_{00}} a_i^{00} = \sum_{j=1}^{s_{00}} b_j^{00}$. Ostatnie warunki zapewniają, że $\ell_1(x) = \ell_2(x)$. Zostawiamy te

same założenia dotyczące współczynników a_i i b_j , co w dowodzie lematu A.2, odwołując się do drugiego lematu o małym zaburzeniu.

Zapisujemy $x = x_{00} + x_{01} + x_{10}$, gdzie $x_{00} = \sum_{i=1}^{r_{00}} a_i^{00} e_i^{00} + \sum_{j=1}^{s_{00}} b_j^{00} f_j^{00}$. Podobnie w przypadku x_{01} i x_{10} . Dla każdego z trzech składników tej sumy przeprowadzamy konstrukcję analogiczną jak w dowodzie lematu A.2. Na przykład w przypadku x_{00} dostajemy przedstawienie za pomocą $r_{00} + s_{00} - 1$ składników takich, że wszystkie współczynniki są dodatnie i sumują się do $\sum_{i=1}^{r_{00}} a_i^{00}$. Możemy zapisać

$$x = \sum c_{i,j}^{00}(e_i^{00} + f_j^{00}) + \sum c_{i,j}^{01}(e_i^{01} + f_j^{01}) + \sum c_{i,j}^{10}(e_i^{10} + f_j^{10}),$$

gdzie z konstrukcji wynika, że $\sum c_{i,j}^{00} + \sum c_{i,j}^{01} + \sum c_{i,j}^{10} = 1$.

Pozostaje pokazać, że tak uzyskane zbiory wektorów

$$\{e_i^{00} + f_j^{00} : c_{i,j}^{00} \neq 0, (i, j) \neq (1, 1)\}, \{e_i^{01} + f_j^{01} : c_{i,j}^{01} \neq 0\}, \{e_i^{10} + f_j^{10} : c_{i,j}^{10} \neq 0\}$$

po odjęciu od nich wektora $e_1^{00} + f_1^{00}$ utworzą bazę dla produktu włóknistego krat $M = M_1 \times_{\ell_1 \times \ell_2} M_2$. Oznaczmy ten zbiór \mathcal{B} , to znaczy

$$\mathcal{B} = \{(e_i^{00} + f_j^{00}) - (e_1^{00} + f_1^{00}), (e_i^{01} + f_j^{01}) - (e_1^{00} + f_1^{00}), (e_i^{10} + f_j^{10}) - (e_1^{00} + f_1^{00}) : i, j \text{ jak wyżej}\}.$$

Na mocy założenia unimodularności mamy, że zbiór wektorów złożony z trzech bloków

$$\{e_2^{00} - e_1^{00}, \dots, e_{r_{00}}^{00} - e_1^{00}; e_1^{01} - e_1^{00}, \dots, e_{r_{01}}^{01} - e_1^{00}; e_1^{10} - e_1^{00}, \dots, e_{r_{10}}^{10} - e_1^{00}\}$$

tworzy bazę kraty M_1 . Możemy go łatwo zamienić (odejmując w drugim bloku wektor $e_1^{01} - e_1^{00}$, a w trzecim $e_1^{10} - e_1^{00}$) na równoważny układ bazowy

$$\{e_2^{00} - e_1^{00}, \dots, e_{r_{00}}^{00} - e_1^{00}; \mathbf{e}_1^{01} - \mathbf{e}_1^{00}, e_2^{01} - e_1^{01}, \dots, e_{r_{01}}^{01} - e_1^{01}; \mathbf{e}_1^{10} - \mathbf{e}_1^{00}, e_2^{10} - e_1^{10}, \dots, e_{r_{10}}^{10} - e_1^{10}\}.$$

Podobnie w przypadku kraty M_2 możemy pokazać, że jest ona minimalnie generowana przez zbiór

$$\{f_2^{00} - f_1^{00}, \dots, f_{s_{00}}^{00} - f_1^{00}; \mathbf{f}_1^{01} - \mathbf{f}_1^{00}, f_2^{01} - f_1^{01}, \dots, f_{s_{01}}^{01} - f_1^{01}; \mathbf{f}_1^{10} - \mathbf{f}_1^{00}, f_2^{10} - f_1^{10}, \dots, f_{s_{10}}^{10} - f_1^{10}\}.$$

W obu przypadkach wytłuszczonym drukiem wyróżniliśmy wektory spoza jądra odpowiednio ℓ_1 i ℓ_2 . Takie reprezentacje baz obu krat będą wygodne, gdyż teraz możemy zauważyć, że krata $M = M_1 \times_{\ell_1 \times \ell_2} M_2 = (M_1 \times M_2) \cap \ker(\ell_1 - \ell_2)$ generowana jest przez wektory

$$\begin{aligned} & \{(e_i^{kl} - e_1^{kl}) : i \geq 2\} \\ & \{(f_i^{kl} - f_1^{kl}) : i \geq 2\} \text{ oraz} \\ & (e_1^{01} + f_1^{01}) - (e_1^{00} + f_1^{00}), \quad (e_1^{10} + f_1^{10}) - (e_1^{00} + f_1^{00}) \end{aligned} \quad (\text{A.1})$$

dla wszystkich możliwych $(k, l) \in \{(0, 0), (0, 1), (1, 0)\}$. Istotnie. Dwa pierwsze zbiory są jasne, bo reprezentują generatory produktu, które leżą w jądrze homomorfizmu krat $\ell_1 - \ell_2$. Dwa ostatnie wektory znajdują się w zbiorze generatorów gdyż kombinacja

$$a(e_1^{01} - e_1^{00}) + b(e_1^{10} - e_1^{00}) + c(f_1^{01} - f_1^{00}) + d(f_1^{10} - f_1^{00})$$

leży w $\ker(\ell_1 - \ell_2)$ wtedy i tylko wtedy, gdy $a = c$ i $b = d$, a zatem powyższą kombinację możemy zamienić na

$$a((e_1^{01} + f_1^{01}) - (e_1^{00} + f_1^{00})) + b((e_1^{10} + f_1^{10}) - (e_1^{00} + f_1^{00})).$$

Wszystkie z wymienionych wektorów dają się przedstawić za pomocą wektorów ze zbioru \mathcal{B} . Zauważmy wpraw, że zbiór \mathcal{B} generuje tą samą kratę co zbiór

$$\begin{aligned} \mathcal{B}' = & \{(e_i^{00} + f_j^{00}) - (e_1^{00} + f_1^{00}) : (i, j) \neq (1, 1)\} \cup \\ & \cup \{(e_i^{01} + f_j^{01}) - (e_1^{01} + f_1^{01}) : (i, j) \neq (1, 1)\} \cup \\ & \cup \{(e_i^{10} + f_j^{10}) - (e_1^{10} + f_1^{10}) : (i, j) \neq (1, 1)\} \cup \\ & \cup \{(e_1^{01} + f_1^{01}) - (e_1^{00} + f_1^{00}), (e_1^{10} + f_1^{10}) - (e_1^{00} + f_1^{00})\}, \end{aligned} \quad (\text{A.2})$$

dla wszystkich (i, j) takich jak w zbiorze \mathcal{B} . Oba zbiory są równoliczne, a zatem \mathcal{B} jest bazą M wtedy i tylko wtedy, gdy \mathcal{B}' jest bazą M . Procedura przejścia od \mathcal{B} do \mathcal{B}' jest podobna jak procedura znajdowania równoważnych baz dla M_1 i M_2 , którą przedstawiliśmy przed chwilą.

Zauważmy teraz, że w kracie generowanej przez zbiór \mathcal{B}' leżą wektory $\{(e_i^{kl} - e_1^{kl}) : i \geq 2\}$ i $\{(f_i^{kl} - f_1^{kl}) : i \geq 2\}$, co wynika bezpośrednio z procedury przedstawionej w dowodzie lematu A.2 przeprowadzonej dla każdego z pierwszych składników sumy zbiorów w równaniu A.2. Dwa pozostałe wektory z bazy kraty M (patrz równanie A.1) po prostu leżą w \mathcal{B}' , co wynika z tego, że zarówno $c_{1,1}^{01} \neq 0$ oraz $c_{1,1}^{10} \neq 0$. Równoliczność zbioru \mathcal{B} z bazą kraty M zapewnia, że \mathcal{B} stanowi minimalny zbiór generatorów. Wektorów w bazie kraty M jest

$$\sum_{(k,l)} (r_{kl} - 1) + \sum_{(k,l)} (s_{kl} - 1) + 2 = \sum_{(k,l)} (r_{kl} + s_{kl}) - 4,$$

czyli tyle samo co wektorów w zbiorze \mathcal{B} . Zapewnia to, że \mathcal{B} w istocie stanowi bazę kraty M , co kończy dowód. \square

Uwaga A.7 *W ten sposób pokazaliśmy, że istotnie musi istnieć pokrycie wielościanu $\Delta_1 \ell_1 \times \ell_2$ Δ_2 simpleksami, których wierzchołki leżą w kracie M i ją generują, a zatem są unimodularne. W szczególności pokazaliśmy, że wierzchołki wielościanu $\Delta_1 \ell_1 \times \ell_2$ Δ_2 istotnie leżą w kracie M , a zatem jest to wielościan kratowy.*

Wniosek A.8 *Niech $Q_1 \subset (M_1)_{\mathbb{R}}$ i $Q_2 \subset (M_2)_{\mathbb{R}}$ będą dwoma wielościanami, które posiadają pokrycie unimodularne. Rozważmy dwa homomorfizmy kratowe $\ell_i : M_i \rightarrow \mathbb{Z}^k$ takie, że $\ell_i(Q_i) = \Delta$, gdzie Δ jest standardowym k -wymiarowym simpleksem kratowym. Wówczas produkt włóknisty $Q = Q_1 \ell_1 \times \ell_2 Q_2$ ma pokrycie unimodularne względem produktu włóknistego krat $M = M_1 \ell_1 \times \ell_2 M_2$.*

Dowód: Produkt włóknisty Q ma pokrycie produktami włóknistymi simpleksów z pokryć $Q_1 \subset (M_1)_{\mathbb{R}}$ i $Q_2 \subset (M_2)_{\mathbb{R}}$. Na mocy lematu A.6, jeżeli pokrycia Q_1 i Q_2 są unimodularne odpowiednio względem krat M_1 i M_2 to Q ma pokrycie unimodularne względem kraty M . \square

Dodatek B

Obliczenia

B.1. Wielościany P_3 i R_3

Wielościan P_n ma 2^n wierzchołków postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \cdots \oplus e_{i_{n-1} i_n}^{(n-1)} \oplus e_{i_n i_1}^{(n)},$$

gdzie $(i_1, \dots, i_n) \in \{0, 1\}^n$ wektor $e_{ij}^{(k)}$ dla każdej z n kopii przestrzeni \mathbb{R}^4 (indeksowanych superskrytem k) jest wektorem jednostkowym tej przestrzeni. Przyjmujemy konwencję

$$e_{00}^{(k)} = [1, 0, 0, 0], e_{01}^{(k)} = [0, 1, 0, 0], e_{10}^{(k)} = [0, 0, 1, 0], e_{11}^{(k)} = [0, 0, 0, 1]$$

dla każdego $k = 1, \dots, n$.

Poniższa tabela zawiera współrzędne wielościanu P_3 . Każdy wiersz odpowiada pewnemu wierzchołkowi ($2^3 = 8$).

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Wielościan R_n odpowiadający modelowi liniowemu ma 2^n wierzchołków postaci

$$e_{i_1 i_2}^{(1)} \oplus e_{i_2 i_3}^{(2)} \oplus \cdots \oplus e_{i_{n-1} i_n}^{(n-1)},$$

gdzie $(i_1, \dots, i_n) \in \{0, 1\}^n$ z oznaczeniami jak w przypadku wielościanu P_n . Z samej konstrukcji wierzchołków widać, że wielościan R_n uzyskamy z P_n zapominając o czterech ostatnich współrzędnych (jak pokazały obliczenia, wymiar wielościanu spadnie o jeden). W związku z tym łatwo przedstawić tabelę ze współrzędnymi wierzchołków wielościanu R_3 .

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

B.2. Obliczenia w Polymake

Program `Polymake` został stworzony do obliczeń dla wymiernych wielościanów wypukłych. Obecnie jednak jego funkcjonalności została znacznie rozszerzona. Szczegółowe informacje można znaleźć na stronie: <http://www.math.tu-berlin.de/polymake/>.

W naszym przypadku główne zadania wykonane za pomocą `Polymake` to znajdowanie triangulacji wielościanów kratowych, badanie kombinatoryki tych wielościanów (liczba ścian poszczególnych wymiarów) oraz badanie generatorów poszczególnych stożków wachlarzy normalnych. Większość obliczeń była dokonywana dla wielościanów uzyskanych po włożeniu w podprzestrzeń afiniczną w której były one już pełnowymiarowe. Poniżej zamieszczamy wyniki kluczowych obliczeń dla wielościanu P_n w przypadku $n = 3, 4$. Analiza triangulacji dla wielościanów P_n dla $n \leq 7$ pod względem unimodularności została dokonana przy pomocy skryptu zamieszczonego w części B.3.

Krótką legenda do wyników: Poniżej `VERTICES` znajdują się współrzędne poszczególnych wierzchołków (poprzedzone jedynką), `TRIANGULATION` to zbiory wierzchołków definiujących poszczególne simpleksy triangulacji (numeracja wierzchołków od zera zależnie od kolejności poniżej `VERTICES`). `FACETS` to współrzędne w przestrzeni dualnej nierówności definiujących nierówności podpierające ściany maksymalnego wymiaru. `VERTICES_IN_FACETS` to zbiory wierzchołków leżących na poszczególnych ścianach maksymalnego wymiaru. `F_VECTOR` to wektor, którego elementy mówią o liczbie ścian danego wymiaru (od zera do $\dim P - 1$), gdzie pierwszy jego element to liczba wierzchołków, drugi to liczba krawędzi, a ostatni to liczba ścian maksymalnego wymiaru wielościanu P . Ostatecznie `F2_VECTOR` to macierz incydencji dla ścian wielościanu, gdzie element f_{ij} poza diagonalą oznacza liczbę sąsiadujących par ścian wymiaru i i ścian wymiaru j . Diagonala to `F_VECTOR`.

B.2.1. Wielościan P_3

VERTICES	FACETS	VERTICES_IN_FACETS	F_VECTOR
1 0 0 0 0 0 0	0 0 0 0 -1 1 1	{0 2 3 4 6 7}	8 28 56 68 48 16
1 0 0 0 0 1 0	0 0 -1 1 0 0 1	{0 1 3 4 6 7}	
1 0 0 1 0 0 0	0 0 0 0 0 0 1	{0 1 2 3 4 6}	F2_VECTOR
1 0 0 0 1 1 0	0 1 0 0 0 0 0	{0 1 2 3 6 7}	8 56 168 272 240 96
1 1 0 0 0 0 0	1 0 0 -1 0 -1 -1	{1 2 3 5 6 7}	56 28 168 408 480 240
1 1 0 0 0 0 1	1 -1 0 -1 0 -1 0	{1 2 3 4 5 6}	168 168 56 272 480 320
1 0 1 1 0 0 0	1 -1 -1 0 0 -1 0	{1 3 4 5 6 7}	272 408 272 68 240 240
1 0 1 0 1 0 1	1 -1 0 -1 -1 0 0	{2 3 4 5 6 7}	240 480 480 240 48 96
	0 1 0 0 1 0 -1	{0 1 2 5 6 7}	96 240 320 240 96 16
TRIANGULATION	0 1 1 0 0 0 -1	{0 1 2 3 5 7}	
{0 1 2 3 4 5 6}	0 0 0 0 1 0 0	{0 1 2 4 5 6}	
{0 2 3 4 5 6 7}	0 0 1 0 0 0 0	{0 1 2 3 4 5}	
{0 1 3 4 5 6 7}	0 0 -1 1 1 0 0	{0 1 4 5 6 7}	
{0 1 2 3 5 6 7}	0 0 0 1 0 0 0	{0 1 3 4 5 7}	
	0 0 0 0 0 1 0	{0 2 4 5 6 7}	
	0 0 1 0 -1 1 0	{0 2 3 4 5 7}	

B.2.2. Wielościan P_4

POINTS	VERTICES_IN_FACETS	TRIANGULATION
1 0 0 0 0 0 0 0 0	{0 1 3 4 5 7 8 9 11 12 13 15}	{0 1 2 3 4 6 8 9 12}
1 0 0 0 0 0 0 0 1 0	{0 1 2 3 4 5 6 7 9 11 13 15}	{1 2 3 4 5 6 8 9 12}
1 0 0 0 0 1 0 0 0 0	{0 4 6 7 8 9 11 15}	{0 1 3 4 6 7 8 9 12}
1 0 0 0 0 0 1 1 0	{0 2 3 4 11 12 13 15}	{1 3 4 5 6 7 8 9 12}
1 0 0 1 0 0 0 0 0 0	{0 1 2 3 4 5 6 7 12 13 14 15}	{2 3 4 5 6 8 9 10 12}
1 0 0 1 0 0 0 1 0	{0 1 3 7 8 12 14 15}	{3 4 5 6 7 8 9 10 12}
1 0 0 0 1 1 0 0 0 0	{0 1 2 6 9 13 14 15}	{1 2 3 5 6 8 9 10 12}
1 0 0 0 1 0 1 1 0	{0 1 4 5 6 7 8 9 12 13 14 15}	{1 3 5 6 7 8 9 10 12}
1 1 0 0 0 0 0 0 0 0	{1 2 3 5 10 12 13 14}	{0 2 3 4 6 8 9 11 12}
1 1 0 0 0 0 0 0 0 1	{1 5 6 7 8 9 10 14}	{0 3 4 6 7 8 9 11 12}
1 1 0 0 0 1 0 0 0 0	{3 4 5 7 8 10 11 12}	{3 4 6 7 8 9 10 11 12}
1 1 0 0 0 0 1 0 0 1	{2 4 5 6 9 10 11 13}	{2 3 4 6 8 9 10 11 12}
1 0 1 1 0 0 0 0 0 0	{0 1 2 4 5 6 8 9 10 12 13 14}	{3 4 5 6 7 9 10 11 12}
1 0 1 1 0 0 0 0 0 1	{0 1 2 3 4 5 6 7 8 10 12 14}	{2 3 4 5 6 9 10 11 12}
1 0 1 0 1 1 0 0 0 0	{2 3 4 5 6 7 10 11 12 13 14 15}	{3 4 5 7 8 9 10 11 12}
1 0 1 0 1 0 1 0 0 1	{4 5 6 7 8 9 10 11 12 13 14 15}	{2 4 5 6 9 10 11 12 13}
	{0 2 3 4 6 7 8 10 11 12 14 15}	{1 3 4 5 6 7 9 12 13}
	{0 1 2 3 4 5 6 7 8 9 10 11}	{0 1 3 4 6 7 9 12 13}
	{1 2 3 5 6 7 9 10 11 13 14 15}	{1 2 3 4 5 6 9 12 13}
	{0 1 2 3 6 7 8 9 10 11 14 15}	{0 1 2 3 4 6 9 12 13}
	{0 1 2 3 4 5 8 9 10 11 12 13}	{0 3 4 6 7 9 11 12 13}
	{0 2 4 6 8 9 10 11 12 13 14 15}	{0 2 3 4 6 9 11 12 13}
	{0 1 2 3 8 9 10 11 12 13 14 15}	{2 3 4 5 6 9 11 12 13}
	{1 3 5 7 8 9 10 11 12 13 14 15}	{3 4 5 6 7 9 11 12 13}
		{1 3 5 6 7 9 10 12 13}
		{1 2 3 5 6 9 10 12 13}
		{2 3 5 6 9 10 11 12 13}
		{3 5 6 7 9 10 11 12 13}
		{1 2 3 5 6 10 12 13 14}
		{1 3 5 6 7 10 12 13 14}
		{0 1 2 3 6 8 9 12 14}
		{0 1 3 6 7 8 9 12 14}
		{1 3 6 7 8 9 10 12 14}
		{1 2 3 6 8 9 10 12 14}
		{2 3 6 8 9 10 11 12 14}
		{3 6 7 8 9 10 11 12 14}
		{0 3 6 7 8 9 11 12 14}
		{0 2 3 6 8 9 11 12 14}
		{0 2 3 6 9 11 12 13 14}
		{0 3 6 7 9 11 12 13 14}
		{0 1 2 3 6 9 12 13 14}
		{0 1 3 6 7 9 12 13 14}
		{3 6 7 9 10 11 12 13 14}
		{2 3 6 9 10 11 12 13 14}
		{1 2 3 6 9 10 12 13 14}
		{1 3 6 7 9 10 12 13 14}
		{1 5 6 7 8 9 10 12 14}
		{1 5 6 7 9 10 12 13 14}
		{0 2 3 4 6 11 12 13 15}
		{0 3 4 6 7 11 12 13 15}
		{0 3 6 7 11 12 13 14 15}
		{0 2 3 6 11 12 13 14 15}
		{0 1 3 7 9 12 13 14 15}
		{0 3 7 9 11 12 13 14 15}
		{0 3 7 8 9 11 12 14 15}
		{0 1 3 7 8 9 12 14 15}
		{0 1 3 6 7 9 13 14 15}
		{0 1 2 3 6 9 13 14 15}
		{0 3 6 7 9 11 13 14 15}
		{0 2 3 6 9 11 13 14 15}
		{0 4 6 7 8 9 11 12 15}
		{0 4 6 7 9 11 12 13 15}
		{0 6 7 9 11 12 13 14 15}
		{0 6 7 8 9 11 12 14 15}

FACETS	F_VECTOR	F2_VECTOR
0 0 0 0 0 1 0 0 0 0	16 104 360 712 816 520 168 24	16 208 1088 2912 4272 3392 1376 256
0 0 1 1 0 0 0 0 0 -1		208 104 1088 4400 8752 8960 4672 1152
0 0 1 0 -1 1 0 1 0		1088 1088 360 2912 8768 12240 8480 2816
0 1 0 0 1 0 -1 1 0		2912 4400 2912 712 4288 9088 8832 4032
0 1 0 0 0 0 0 0 0		4272 8752 8768 4288 816 3456 5216 3392
0 0 -1 1 0 1 0 0 1		3392 8960 12240 9088 3456 520 1568 1568
0 1 0 1 0 0 1 0 -1		1376 4672 8480 8832 5216 1568 168 336
0 0 0 0 -1 1 1 0 0		256 1152 2816 4032 3392 1568 336 24
1 0 -1 0 1 -1 0 -1 0		
1 -1 0 0 -1 0 1 -1 0		
1 -1 0 -1 0 0 -1 0 1		
1 0 1 -1 0 -1 0 0 -1		
0 0 0 0 0 0 1 0 0		
0 0 0 0 0 0 0 0 1		
1 0 0 -1 0 -1 -1 0 0		
1 -1 0 -1 -1 0 0 0 0		
0 0 0 0 0 -1 1 1		
0 0 1 0 0 0 0 0 0		
1 0 0 0 0 -1 0 -1 -1		
0 0 0 1 0 0 0 0 0		
0 0 0 0 1 0 0 0 0		
0 0 0 0 0 0 0 1 0		
0 0 -1 1 1 0 0 0 0		
1 -1 -1 0 0 0 0 -1 0		

B.3. Użyteczny skrypt dla Polymake

Zakładamy, że posiadamy plik, w którym Polymake przedstawił triangulację zadanego wielościanu. W naszych zastosowaniach interesowało nas, czy unormowana objętość wszystkich simpleksów w tej triangulacji wynosi jeden (czy triangulacja jest unimodularna). W Polymake nie istnieje opcja dzięki, której można automatycznie wykonać to zadanie, a z drugiej strony sprawdzanie każdego simpleksu triangulacji osobno staje się niezwykle uciążliwe.

Poniższy skrypt został stworzony przez Michaela Joswiga, któremu autor serdecznie dziękuje za okazaną pomoc. Skrypt ten pozwala na automatyczne wyliczenie wszystkich objętości simpleksów w triangulacji. Przedstawiamy go tutaj, gdyż nigdzie dotąd nie został opublikowany.

W pierwszej kolejności wpisujemy do pustego pliku:

```
application 'polytope';

die "usage: polymake --script triangulation_volumes FILE\n" unless @ARGV;

my $p=load($ARGV[0]);
my @vertices= @{$p->VERTICES};

foreach my $this_face (@{$p->TRIANGULATION}) {
    my @this_simplex= $this_face =~ /(\d+)/g;
    my @these_vertices= @vertices[@this_simplex];
    my $simplex=new Apps::polytope::RationalPolytope("some simplex",
VERTICES=>\@these_vertices);
    print "The simplex {@this_simplex} has volume ", $simplex->VOLUME, ".\n";
}
```

Teraz możemy go zapisać na przykład pod nazwą `triangulation_volumes` w katalogu, w którym Polymake przechowuje skrypty. Niech `FILE` będzie plikiem zawierającym dane wielościanu i jego triangulacji. Wpisując wówczas komendę

```
polymake --script triangulation_volumes FILE
```

uzyskujemy na wyjściu objętość każdego z simpleksów triangulacji.

Bibliografia

- [Agr02] Alan Agresti. *Categorical Data Analysis*. John Wiley & Sons, New York, wydanie drugie, 2002.
- [BGT97] Winfried Bruns, Joseph Gubeladze, Ngo Viet Trung. Normal polytopes, triangulations, and Koszul algebras. *J. Reine Angew. Math*, 485:123–160, 1997.
- [BH98] Winfried Bruns, Jürgen Herzog. *Cohen-Macaulay Rings*. Cambridge University Press, 1998.
- [BW06] Weronika Buczyńska, Jarosław A. Wiśniewski. On phylogenetic trees – a geometer’s view. [arXiv:math.AG/0601357](https://arxiv.org/abs/math/0601357), 2006.
- [Cox] David Cox. Lectures on toric varieties. <http://www.cs.amherst.edu/~dac/lectures/coxcimpa.pdf>.
- [DS98] Persi Diaconis, Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, 26(1):363–397, 1998.
- [DS03] Mike Develin, Seth Sullivant. Markov Bases of Binary Graph Models. *Annals of Combinatorics*, 7(4):441–466, 2003.
- [DS06] Mathias Drton, Seth Sullivant. Algebraic Statistical Models. http://www.math.harvard.edu/~seths/ds10_23final.pdf, 2006.
- [Edw00] David Edwards. *Introduction to graphical modelling*. Springer, wydanie second, 2000.
- [ERSS04] Nicholas Eriksson, Kristian Ranestad, Bernd Sturmfels, Seth Sullivant. Phylogenetic Algebraic Geometry. [arXiv:math.AG/0407033](https://arxiv.org/abs/math/0407033), 2004.
- [Ful93] William Fulton. *Introduction to toric varieties*. Princeton University Press Princeton, NJ, 1993.
- [GJ05] Ewgenij Gawrilow, Michael Joswig. Geometric reasoning with polymake. [arXiv:math.CO/0507273](https://arxiv.org/abs/math/0507273), 2005.
- [GMS02] Dan Geiger, Christopher Meek, Bernd Sturmfels. On the toric algebra of graphical models, 2002.
- [GP93] Dan Geiger, Judea Pearl. Logical and Algorithmic Properties of Conditional Independence and Graphical Models. *The Annals of Statistics*, 21(4):2001–2021, 1993.

- [Haa] Christian Haase. Lattice polytopes and triangulations with applications to toric geometry. <http://citeseer.ist.psu.edu/626167.html>.
- [Har77] Robin Hartshorne. *Algebraic Geometry*. Springer, 1977.
- [HHM05] Ralf Hemmecke, Raymond Hemmecke, Peter Malkin. 4ti2 version 1.2—computation of Hilbert bases, Graver bases, toric Gröbner bases, and more. <http://www.4ti2.de/>, 2005.
- [HS02] Serkan Hoşten, Seth Sullivant. Grobner bases and polyhedral geometry of reducible and cyclic models. *J. Combin. Theory Ser. A*, 100(2):277–301, 2002.
- [KKMSD73] George R. Kempf, Finn F. Knudsen, David Mumford, Bernard Saint-Donat. *Toroidal embeddings I, LNM 339*, wolumen 339 serii *Springer Lecture Notes*. Springer-Verlag, 1973.
- [Lau96] Steffen L. Lauritzen. *Graphical models*. Number 17 serii Oxford Statistical Science Series. Clarendon Press, Oxford, 1996.
- [PRW01] Giovanni Pistone, Eva Riccomagno, Henry P. Wynn. *Algebraic Statistics: Computational Commutative Algebra in Statistics*. Chapman&Hall, 2001.
- [SP05] Bernd Sturmfels, Lior Patcher, redaktorzy. *Algebraic Statistics for Computational Biology*. Cambridge University Press, 2005.
- [SS06] Bernd Sturmfels, Seth Sullivant. Toric geometry of cuts and splits. [arXiv: math.AC/0606683](http://arxiv.org/abs/math.AC/0606683), 2006.
- [Stu96] Bernd Sturmfels. *Gröbner bases and convex polytopes*, wolumen 8 serii *University Lecture Series*. American Mathematical Society, Providence, RI, 1996.
- [Stu06] Bernd Sturmfels. Can Biology Lead To New Theorems? <http://math.berkeley.edu/~bernd/ClayBiology.pdf>, 2006.
- [Sul05] Seth Sullivant. *Toric Ideals in Algebraic Statistics*. Praca doktorska, University of California, Berkeley, 2005.
- [Whi90] Joe Whittaker. *Graphical Models in Applied Multivariate Statistics*. Wiley New York, 1990.
- [Wil04] Robin J. Wilson. *Wprowadzenie do teorii grafów*. Państwowe Wydawnictwo Naukowe, wydanie II, 2004.
- [Zie90] Ryszard Zieliński. *Siedem wykładów wprowadzających do statystyki matematycznej*. PWN, Warszawa, 1990.