Kapitel 3
Anwendung auf Designs

Eine der Grundaufgaben der Statistik besteht darin, ausgehend von Stichproben Modelle zu entwickeln und deren Parameter zu schätzen. Wir gehen aus von einem Körper k, den wir meist als Teilkörper der reellen Zahlen auffassen werden. Da wir allerdings im Körper k exakt rechnen wollen, sollte k so klein wie möglich sein, etwa $k = \mathbb{Q}$.

Definition: Ein *Design* D in k^n ist eine endliche Teilmenge von k^n.

Oftmals haben die n Koordinaten der Punkte aus D inhaltliche Interpretationen, indem sie die Ausprägungen verschiedener Faktoren kodieren. Eine Stichprobe dient dann dazu, den Einfluß verschiedener der Faktoren auf ein Ergebnis abzuschätzen. Beispiele sind etwa der Ertrag eines landwirtschaftlichen Produkts in Abhängigkeit von Düngung, Bewässerung, Bodenbeschaffenheit *usw.,* oder der Preis, den ein Kunde für ein Auto zu zahlen bereit ist in Abhängigkeit von verschiedenen Ausstattungsmerkmalen. Oft werden für diese Faktoren nur endlich viele Stufen betrachtet. Diese können auf einer reinen Nominalskala liegen, etwa $\{\textit{rot}, \textit{blau}, \textit{gelb}\}$ oder $\{\textit{vorhanden}, \textit{nicht vorhanden}\}$, weshalb man sich in der Literatur im Falle von ℓ Stufen oft mit der Kodierung durch die Zahlen von Null bis $\ell - 1$ begnügt. Wenn es die Darstellung erleichtert, werde ich dies im Folgenden auch gelegentlich tun, aber inhaltlich ändert sich dadurch nichts.

Definition: *a)* Ein volles faktorielles Design für n Faktoren ist ein kartesisches Produkt von n endlichen Teilmengen $S_i \subset k$. Falls alle M_i jeweils ℓ Elemente haben, sprechen wir von einem ℓ^n-Design.

b) Ein (fraktionelles) faktorielles Design für n Faktoren ist eine Teilmenge eines vollen faktoriellen Designs für diese Faktoren.

Offensichtlich kann jedes Design $D \subset k^n$ als fraktionelles faktorielles Design angesehen werden, etwa für S_i gleich Menge aller möglicher Werte, die die Variable x_i auf D annehmen kann.

§1: Allgemeine lineare Modelle

Wir haben uns bereits in Kapitel 0 mit statistischen Modellen beschäftigt; hier seien noch einmal kurz die im folgenden relevanten bezeichnungen rekapituliert.

Wir starten mit einem Design $D = \{x^{(1)}, \dots, x^{(r)}\} \subset k^n$. Typischerweise ist für jeden Punkt $x^{(j)} \in D$ ein Wert $y_j \in k$ gegeben; gesucht ist eine Funktion $f\colon k^n \to k$ mit $f(x^{(j)}) = y_j$ oder zumindest $f(x^{(j)}) \approx y_j$ für alle j.

Für den allgemeinsten (linearen) Ansatz zum Auffinden solcher Funktionen wählen wir eine endliche Menge $\mathbb{F} = \{f_1, \dots, f_s\}$ von Funktionen $f_i\colon k^n \to k$ und betrachten Funktionen f der Form $f = \sum_{i=1}^s \theta_i f_i$ mit $\theta_1, \dots, \theta_s \in k$.

Definition: Die Z-Matrix zu $\mathbb{F}$ und D ist die $s \times r$-Matrix Z mit Einträgen $z_{ij} = f_i(x^{(j)})$.

Für $f = \sum\limits_{i=1}^{s} \theta_i f_i$ sollte dann idealerweise gelten

$$y_j = \sum_{i=1}^{s} \theta_i f_i(x^{(j)}) = \sum_{i=1}^{s} \theta_i z_{ij}\,;$$

ausgedrückt mit den Vektoren

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix} \quad \text{und} \quad \theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_s \end{pmatrix}$$

wird dies zu $y^T = \theta^T Z \quad \text{oder} \quad Z^T \theta = y\,.$

Falls $r < s$ ist, kann dieses Gleichungssystem keine eindeutig bestimmte Lösung haben. Für $r = s$ gibt es genau dann eine, wenn die Z-Matrix invertierbar ist; dann ist $\theta = (Z^T)^{-1} y$.

In der Statistik ist meist $r > s$; in diesem Fall wird es im allgemeinen keine Lösung geben. Dann müssen wir uns begnügen mit einem Vektor θ derart, daß der Fehler $\varepsilon = y - Z^T\theta$ möglichst klein wird. Wie wir in Kapitel 0 gesehen haben, führt dies auf das lineare Gleichungssystem

$$ZZ^T\theta = Zy\,.$$

ZZ^T ist eine quadratische Matrix; es gibt daher genau dann eine eindeutig bestimmte Lösung, wenn diese Matrix nichtsingulär ist. Das wiederum ist äquivalent dazu, daß Z (und damit auch Z^T) den Rang s hat: In diesem Fall ist nämlich die Abbildung von k^r nach k^s, die einem Vektor x den Vektor Zx zuordnet, surjektiv, und die Abbildung von k^s nach k^r, die θ auf $Z^T\theta$ abbildet, hat ein s-dimensionales Bild. Somit hat die Abbildung von k^r nach k^r, die einem Vektor $x \in k^r$ den Vektor $ZZ^Tx = Z(Z^Tx)$ zuordnet, ein s-dimensionales Bild, d.h. der Rang von ZZ^T ist s. Ist umgekehrt der Rang von Z kleiner als s, so erst recht der von ZZ^T, so daß ZZ^T in der Tat genau dann nichtsingulär ist, wenn der Rang von Z gleich s ist.

Damit haben wir gezeigt

Satz: Falls der Rang von Z gleich s ist, gibt es genau einen Vektor $\theta \in k^s$, für den die Differenz $Z^T\theta - y$ minimale Länge hat, nämlich $\theta = (ZZ^T)^{-1}y$.

■

§2: Polynomiale lineare Modelle

In dieser Vorlesung soll es in erster Linie um Modelle gehen, bei denen die Funktionen aus $\mathbb{F}$ Monome sind. Schon lange vor dem Aufkommen der algebraischen Statistik betrachtete man im Sinne möglichst einfacher Modelle vorzugsweise Mengen $\mathbb{F}$, die mit jedem Monom auch dessen sämtliche Teiler enthalten; diese hatten wir in Kapitel 0 als Ordnungsideale bezeichnet. Um auch andere Modelle nicht ganz auszuschließen, fassen wir die folgende Definition etwas allgemeiner:

Definition: k sei ein Körper, $R = k[X_1, \ldots, X_n]$ ein Polynomring über k, und $\mathbb{T}$ sei die Menge aller Monome $X_1^{e_1} \cdots X_n^{e_n}$ mit $e_i \in \mathbb{N}_0$.
a) Ein lineares Modell mit $\mathbb{F} \subset \mathbb{T}$ heißt *monomiales lineares Modell;*

die Menge $\mathbb{F}$ wird als der *Träger* des Modells bezeichnet.
b) Eine nichtleere Teilmenge $\mathcal{O}$ von $\mathbb{T}$ heißt *Ordnungsideal,* wenn für jedes Monom aus $\mathcal{O}$ auch dessen sämtliche Teiler in $\mathcal{O}$ liegen.
c) Das Modell heißt *vollständig,* wenn $\mathbb{F}$ ein Ordnungsideal ist.

Das Ergebnis am Ende des letzten Kapitels zeigt, wie wir zu jedem Design D ein vollständiges lineares Modell finden können: Wir fassen D zunächst auf als Nullstellenmenge eines Ideals I des Polynomrings R. Wie man ein solches Ideal I finden kann, haben wir am Ende von §1 des vorigen Kapitels gesehen, und ganz am Ende des Kapitels haben wir uns überlegt, daß es das volle Designideal $I(D)$ ist. Danach wählen wir eine Monomordnung τ auf R und berechnen dazu eine GRÖBNER-Basis von I.

Definition: *a)* $\mathrm{Est}_\tau(D)$ ist die Menge aller Standardmonome zu einer GRÖBNER-Basis von $I(D)$ bezüglich der Monomordnung τ.
b) Die Menge aller Mengen $\mathrm{Est}_\tau(D)$, wobei τ die sämtlichen Monomordnungen von R durchläuft, heißt der (GRÖBNER-)*Fächer* von D.

Mit unseren bisherigen Mitteln ist die Konstruktion des Fächers meist recht aufwendig, da wir dazu GRÖBNER-Basen des Designideals zu jeder der unendlich vielen möglichen Monomordnungen kennen und damit auch berechnen müssen.

Im Falle eines vollen faktoriellen Design $D = S_1 \times \cdots \times S_n$ haben wir damit allerdings keinerlei Schwierigkeiten: Offensichtlich bilden die Polynome

$$g_i = \prod_{x \in S_i} (X_i - x)$$

ein Erzeugendensystem von $I(D)$, und dieses Erzeugendensystem ist nach dem Kriterium von BUCHBERGER bezüglich jeder Monomordnung eine GRÖBNER-Basis, denn g_i ist ein Polynom nur in X_i, so daß sein führendes Monom eine Potenz von X_i ist. Somit sind die führenden Monome zweier verschiedener g_i teilerfremd. Wie wir bei der Diskussion von BUCHBERGERs Kriterium gesehen haben reduziert daher $S(g_i, g_j)$ für $i \neq j$ bezüglich jeder Monomordnung auf Null.

Damit ist klar, daß die Standardmonome für ein volles faktorielles Design $D = S_1 \times \cdots \times S_n$ bezüglich jeder Monomordnung genau diejenigen Monome $X_1^{e_1} \cdots X_n^{e_n}$ sind, für die jedes e_i kleiner ist als die Elementanzahl der entsprechenden Menge S_i.

§3: Designs mit minimalem Fächer

Der Fächer eines vollständigen faktoriellen Designs besteht aus genau einem Blatt; kleiner kann er nicht werden. Deshalb definieren wir

Definition: Ein Design D hat minimalen Fächer, wenn alle Monomordnungen τ auf die gleiche Menge $\mathrm{Est}_\tau(D)$ führen.

Das ist insbesondere dann der Fall, wenn alle Monomordnungen zur gleichen GRÖBNER-Basis führen; betrachten wir also diesen Fall etwas genauer:

Definition: *a)* Eine Teilmenge $G = \{g_1, \ldots, g_m\}$ eines Ideals I von $k[X_1, \ldots, X_n]$ heißt *universelle* GRÖBNER-Basis, wenn sie bezüglich jeder Monomordnung auf $k[X_1, \ldots, X_n]$ eine GRÖBNER-Basis ist.
b) G heißt *Super-G-Basis* bezüglich einer Monomordnung τ, wenn jede Teilmenge von G bezüglich τ eine GRÖBNER-Basis des von ihr erzeugten Ideals ist.

Die beiden hier definierten Konzepte haben eigentlich nichts miteinander zu tun; bei den Designs, die wir hier betrachten wollen, ist aber ihr Zusammenspiel nützlich. Wir beginnen mit einem Kriterium zur Charakterisierung von Super-G-Basen:

Lemma: $G \subset I$ ist genau dann eine Super-G-Basis, wenn für je zwei Elemente $f, g \in G$ gilt:

$$\mathrm{FM}_\tau\bigl(\mathrm{ggT}(f,g)\bigr) = \mathrm{ggT}\bigl(\mathrm{FM}_\tau(f), \mathrm{FM}_\tau(g)\bigr)\,. \qquad (*)$$

Beweis durch Induktion nach $m = \#G$:

Der Fall $m = 1$ ist trivial, denn dann ist $I = (f)$ ein Hauptideal, d.h. zu jedem $g \in I$ gibt es ein Polynom h, so daß $g = fh$ ist. Damit ist auch

$\mathrm{FM}_\tau(g) = \mathrm{FM}_\tau(f)\,\mathrm{FM}_\tau(h)$, d.h. $\mathrm{FM}_\tau(I)$ wird von $\mathrm{FM}_\tau(f)$ erzeugt, so daß jedes Erzeugende von I eine GRÖBNER-Basis von I definiert.

Für $m = 2$ ist $G = \{f, g\}$, und da $\{f\}$ und $\{g\}$ GRÖBNER-Basen der Hauptideale (f) und (g) sind, ist $\{f, g\}$ genau dann eine Super-G-Basis, wenn es eine GRÖBNER-Basis ist. Wir müssen daher zeigen, daß $\{f, g\}$ genau dann eine GRÖBNER-Basis von I ist, wenn gilt

$$\mathrm{FM}_\tau\big(\mathrm{ggT}(f, g)\big) = \mathrm{ggT}\big(\mathrm{FM}_\tau(f), \mathrm{FM}_\tau(g)\big)\,.$$

Mit $h = \mathrm{ggT}\big(\mathrm{FM}_\tau(f), \mathrm{FM}_\tau(g)\big)$ ist das kgV der führenden Monome gleich $\mathrm{FM}_\tau(f)\,\mathrm{FM}_\tau(g)/h$, also

$$S(f, g) = \frac{\mathrm{FM}_\tau(g)}{\mathrm{FK}_\tau(f)h} f - \frac{\mathrm{FM}_\tau(f)}{\mathrm{FK}_\tau(g)h} g\,.$$

Nach dem Kriterium von BUCHBERGER ist G genau dann eine GRÖBNER-Basis, wenn dieses Polynom modulo $\{f, g\}$ auf Null reduziert, wenn es also Polynome p, q gibt, so daß $S(f, g) = pf + qg$ ist, wobei weder $\mathrm{FM}_\tau(pf)$ noch $\mathrm{FM}_\tau(qg)$ bezüglich τ größer als das führende Monom von $S(f, g)$ sein darf. Mit

$$\widetilde{g} = \frac{\mathrm{FM}_\tau(g)}{\mathrm{FK}_\tau(f)h} - p \quad \text{und} \quad \widetilde{f} = \frac{\mathrm{FM}_\tau(f)}{\mathrm{FK}_\tau(g)h} - q$$

ist dann

$$\widetilde{g}f - \widetilde{f}g = \frac{\mathrm{FM}_\tau(g)}{\mathrm{FK}_\tau(f)h} f - pf - \frac{\mathrm{FM}_\tau(f)}{\mathrm{FK}_\tau(g)h} g + qg = S(f, g) - (fp + qg) = 0\,.$$

In der Definition von $\widetilde{g}$ *bzw.* $\widetilde{f}$ ist der jeweils erste Summand der führende Term, denn wäre ein Term aus p *bzw.* q größer, so wäre das führende Monom von pf *bzw.* qg größer als das von $S(f, g)$.

Da h der ggT der führenden Monome von f und g ist, sind die führenden Monome von $\widetilde{f}$ und $\widetilde{g}$ teilerfremd, und damit sind auch $\widetilde{f}$ und $\widetilde{g}$ teilerfremd. Wegen $\widetilde{g}f = \widetilde{f}g$ muß daher $\widetilde{f}$ ein Teiler von f sein und $\widetilde{g}$ einer von g. Da Polynomringe faktoriell sind, gibt es somit ein Polynom $\widetilde{h}$ derart, daß $f = \widetilde{h}\widetilde{g}$ und $g = \widetilde{h}\widetilde{f}$ ist. Wegen der Teilerfremdheit von $\widetilde{f}$ und $\widetilde{g}$ folgt daraus, daß $\widetilde{h} = \mathrm{ggT}(f, g)$ sein muß, und wenn wir die führenden Monome betrachten, sehen wir, daß h das führende Monom von $\widetilde{h}$ ist. Falls G eine GRÖBNER-Basis ist, gilt also $(*)$.

Für die Umkehrung nehmen wir der Einfachheit halber an, daß f und g jeweils führenden Koeffizienten eins haben; daduch ändert sich weder das Ideal noch das S-Polynom. Ist dann $(*)$ erfüllt und $\widetilde{h} = \operatorname{ggT}(f, g)$, so gibt es teilerfremde Polynome $\widetilde{f}, \widetilde{g}$ derart, daß $f = \widetilde{h} \cdot \widetilde{f}$ und $g = \widetilde{h} \cdot \widetilde{g}$ ist. Wir können annehmen, daß auch diese Polynome alle die Eins als höchsten Koeffizienten haben, und natürlich ist $\widetilde{g}f - \widetilde{f}g = 0$.

Nach $(*)$ ist das führende Monom h von $\widetilde{h}$ der größte gemeinsame Teiler von $\mathrm{FM}_\tau(f)$ und $\mathrm{FM}_\tau(g)$. Somit ist $\mathrm{FM}_\tau\big(\widetilde{f}\big) = \mathrm{FM}_\tau(f)/h$ und $\mathrm{FM}_\tau\big(\widetilde{g}\big) = \mathrm{FM}_\tau(g)/h$. Schreiben wir

$$\widetilde{f} = \frac{\mathrm{FM}_\tau(f)}{h} - q \quad \text{und} \quad \widetilde{g} = \frac{\mathrm{FM}_\tau(g)}{h} - p$$

mit geeigneten Polynomen p und q, so zeigt die gleiche Rechnung wie oben, daß

$$\widetilde{g}f - \widetilde{f}g = S(f, g) - (pf + qg)$$

ist. Da die linke Seite verschwindet, folgt $S(f, g) = pf + qg$, wobei die führenden Monome von pf und gq nicht größer als das von $S(f, g)$ sind. Also reduziert das S-Polynom auf Null, und G ist nach dem Kriterium von BUCHBERGER eine GRÖBNER-Basis.

Damit ist der Induktionsanfang $m = 2$ gezeigt. Sei nun $m > 2$ und $G = \{g_1, \ldots, g_m\}$. Falls G seine Super-G-Basis ist, muß für je zwei Elemente g_i und g_j auch $\{g_i, g_j\}$ eine GRÖBNER-Basis von (g_i, g_j) sein, d.h. nach dem gerade bewiesenen Fall $m = 2$ gilt obige Gleichung.

Gilt umgekehrt obige Gleichung für alle (i, j) und ist G' eine Teilmenge von G, so ist zunächst nach dem Fall $m = 2$ für jedes Paar (i, j) mit $g_i, g_j \in G'$ die Menge $\{g_i, g_j\}$ eine GRÖBNER-Basis des Ideals (g_i, g_j). Somit reduziert $S(g_i, g_j)$ nach dem Kriterium von BUCHBERGER modulo $\{g_i, g_j\}$ und damit erst recht modulo G' auf Null. Da dies für alle Paare (i, j) gilt, zeigt eine nochmalige Anwendung des Kriteriums von BUCHBERGER, daß G' eine GRÖBNER-Basis des von G' erzeugten Ideals ist. Damit ist G eine Super-G-Basis.

■

Polynome, für die dieses Kriterium sehr einfach nachzuweisen ist, sind *Distraktionen* (Zerstreuungen) von Monomen:

Definition: $M = X_1^{e_1} \cdots X_n^{e_n}$ sei ein Monom, und für jedes $i = 1, \dots, n$ sei ein Vektor $a^{(i)} \in k^{\ell_i}$ gegeben, wobei $\ell_i \geq e_i$. Die *Distraktion* von M bezüglich dieser Vektoren ist

$$D(M) = \prod_{i=1}^{n} \prod_{j=1}^{e_i} \bigl(X_i - a_j^{(i)}\bigr) .$$

Falls alle Vektoren $a^{(i)}$ Nullvektoren sind, ist $D(M) = M$.

Für ein volles faktorielles Design $D = S_1 \times \cdots \times S_n$ mit $\ell_i = \#S_i$ wählen wir Vektoren $a^{(i)}$, deren Komponenten (in irgendeiner Reihenfolge) die verschiedenen Elemente von S_i sind, Dann ist

$$D(X_i^{\ell_i}) = \prod_{j=1}^{\ell_i} \bigl(X_i - a_j^{(i)}\bigr) = \prod_{x \in S_i} \bigl(X_i - x\bigr)$$

und

$$I(D) = \bigl(D(X_1^{\ell_1}), \dots, D(X_n^{\ell_n})\bigr) .$$

Die uns bereits bekannte Tatsache, daß dieses Erzeugendensystem eine universelle GRÖBNER-Basis ist, folgt nun auch aus dem folgenden

Satz: $M_1, \dots, M_r$ seien Monome aus $k[X_1, \dots, X_n]$, und $a^{(1)}, \dots, a^{(n)}$ seien Vektoren über k mit hinreichend vielen Komponenten. Dann bilden die Polynome $D(M_1), \dots, D(M_r)$ eine universelle GRÖBNER-Basis des von ihnen erzeugten Ideals. Falls keines der Monome M_i ein anderes teilt, ist diese GRÖBNER-Basis reduziert.

Beweis: Aus der Definition von $D(M)$ folgt sofort, daß alle Monome in $D(M)$ Teiler von M sind; daher ist M bezüglich jeder Monomordnung das führende Monom von $D(M)$. Außerdem ist klar, daß zwei Distraktionspolynome $D(M_i)$ und $D(M_j)$ (bezüglich der gleichen Vektoren) genau dann Teiler voneinander sind, wenn dasselbe für M_i und M_j gilt. Damit ist insbesondere der ggT zweier Polynome $D(M_i)$ und $D(M_j)$ gleich $D\bigl(\mathrm{ggT}(M_i, M_j)\bigr)$. Da das führende Monom einer Distraktion $D(M)$ gleich M ist, ist also das führende Monom von $\mathrm{ggT}\bigl(D(M_i), D(M_j)\bigr)$ gleich dem ggT der führenden Monome von $D(M_i)$ und $D(M_j)$, so daß das Kriterium aus obigem Lemma erfüllt ist.

Somit bilden die $D(M_i)$ bezüglich jeder Monomordnung eine Super-G-Basis des von ihnen erzeugten Ideals, insbesondere also eine GRÖBNER-Basis.

Diese GRÖBNER-Basis ist minimal genau dann, wenn kein führendes Monom eines $D(M_i)$ das eines $D(M_j)$ mit $j \neq i$ teilt, wenn also kein M_i ein anderes teilt. (Der führende Koeffizient der Distraktion eines Monoms ist natürlich immer gleich eins.) In diesem Fall ist die Basis auch reduziert, denn würde M_i irgendein Monom von $D(M_j)$ teilen, so auch M_j selbst, da alle Monome in $D(M_j)$ Teiler von M_j sind.

■

Außer auf den Fall der vollständigen faktoriellen Designs können wir diesen Satz auch auf sogenannte *Stufendesigns* anwenden:

Definition: Ein Design $D \subset \mathbb{N}_0^n$ heißt *Stufendesign,* wenn für jeden Punkt $(x_1, \dots, x_n) \in D$ auch die sämtlichen Punkte $(u_1, \dots, u_n) \in \mathbb{N}_0^n$ mit $u_i \leq x_i$ für alle i in D liegen.

Als Beispiel eines Stufendesigns, das kein volles faktorielles Design ist, können wir etwa die Menge

$$D = \{0, 1, 2\} \times \{0, 1\} \cup \{0, 1\} \times \{0, 1, 2\} = \{0, 1, 2\} \times \{0, 1, 2\} \smallsetminus \{(2, 2\}$$

betrachten. In der graphischen Darstellung sieht man die Stufe:

$$\begin{array}{cccc} 2 & \bullet & \bullet & \\ 1 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ & 0 & 1 & 2 \end{array}$$

Lemma: Jedes Stufendesign ist die Nullstellenmenge eines Ideals, das von Distraktionen von Monomen bezüglich hinreichend langer Vektoren $a^{(i)} = (0, 1, \dots, \ell_i)$ erzeugt wird.

Beweis durch Induktion nach n: Für $n = 1$ gibt es ein $\ell_1 \in \mathbb{N}_0$, so daß

$$D = \{0, 1, \dots, \ell_1\} = V\big(X_1(X_1 - 1) \dots (X_1 - \ell_1)\big) = V\big(D(X_1^{\ell_1+1})\big)$$

ist.

Ist $n > 1$ und kann x_n Werte aus $\{0, \ldots, \ell_n\}$ annehmen, so beachten wir zunächst, daß für jeden dieser Werte a auch

$$\big\{(x_1, \ldots, x_{n-1}) \in \mathbb{N}^{n-1} \bigm| (x_1, \ldots, x_{n-1}, a) \in D\big\}$$

ein Stufendesign ist, also Nullstellenmenge einer Menge $\mathcal{M}_a$ von Distraktionen von Monomen in $X_1, \ldots, X_{n-1}$. Offensichtlich ist dann D die Nullstellenmenge der Polynome $D(MX_n^{a+1})$ mit $M \in \mathcal{M}_a$ für $a = 0, \ldots, \ell_n$ und von $D(X_n^{\ell_n+1})$.

■

Man beachte, daß dieser Beweis konstruktiv ist und daß die Distraktionen nach dem vorigen Satz eine GRÖBNER-Basis des von ihnen erzeugten Ideals bilden. Wir können damit für Stufendesigns ohne den Umweg über den BUCHBERGER-Algorithmus eine GRÖBNER-Basis des zugehörigen Ideals finden. Da diese sogar universell ist, folgt

Korollar: Jedes Stufendesign hat minimalen Fächer.

■

§4: Fraktionen eines vollen faktoriellen Designs

Nun sei $\mathcal{F}$ eine Teilmenge eines vollen faktoriellen Designs D, und G sei die reduzierte universelle GRÖBNER-Basis von $I(D)$. Dann ist $I(D)$ eine Teilmenge von $I(\mathcal{F})$; es gibt also ein Erzeugendensystem von $I(\mathcal{F})$, das G enthält. Wir können versuchen, das zur Berechnung von Modellen zu $\mathcal{F}$ zu verwenden. Ein erstes technisches Hilfsmittel dazu ist das folgende

Lemma: $\mathcal{O} \subset \mathbb{T}$ sei ein Ordnungsideal. Dann gibt es eine eindeutig bestimmte minimale Teilmenge $\mathrm{Min}(\mathcal{O})$ von $\mathbb{T}$ derart, daß $\mathbb{T} \setminus \mathcal{O}$ genau aus den Vielfachen der Monome aus $\mathrm{Min}(\mathcal{O})$ besteht.

Beweis: Das von $\mathbb{T} \setminus \mathcal{O}$ erzeugte monomiale Ideal $I \lhd k[X_1, \ldots, X_n]$ enthält kein Monom aus $\mathcal{O}$, denn wie wir wissen, enthält ein monomiales Ideal genau die Monome, die durch eines der erzeugenden Monome teilbar sind. Da ein Ordnungsideal mit jedem Monom auch dessen sämtliche Teiler enthält, müßte im Falle eines Monoms aus $\mathcal{O}$ in I das Erzeugendensystem $\mathbb{T} \setminus \mathcal{O}$ ein Element von $\mathcal{O}$ enthalten, was natürlich absurd ist.

Nach dem Lemma von DICKSON hat I ein Erzeugendensystem aus endlich vielen Monomen. Ein solches Erzeugendensystem ist minimal, wenn keines der Monome dort durch ein anderes teilbar ist. Ein gegebenes endliches Erzeugendensystem läßt sich problemlos auf ein minimales reduzieren; also gibt es minimale Erzeugendensysteme.

Angenommen, $\{M_1, \ldots, M_p\}$ und $\{N_1, \ldots, N_s\}$ sind zwei solche minimale Erzeugendensysteme. Da jedes N_i im von den M_j erzeugten monomialen Ideal I liegt, muß N_i durch (mindestens) ein M_j teilbar sein. Da I auch von $N_1, \ldots, N_s$ erzeugt wird, muß umgekehrt M_j durch ein N_ℓ teilbar sein, d.h. $N_\ell | M_j | N_i$. Da wir nur minimale Erzeugendensysteme betrachten, folgt $N_\ell = N_i$, also insbesondere $N_i = M_j$. Jedes N_i liegt also im Erzeugendensystem $\{M_1, \ldots, M_p\}$, und da dies ein minimales Erzeugendensystem ist, folgt $\{M_1, \ldots, M_p\} = \{N_1, \ldots, N_q\}$, d.h. es gibt genau ein minimales Erzeugendensystem von I. Da die Monome in I genau die aus $\mathbb{T} \setminus \mathcal{O}$ sind, besteht diese Menge somit genau aus den Vielfachen der Monome aus $\{M_1, \ldots, M_p\}$, so daß wir $\mathrm{Min}(\mathcal{O}) = \{M_1, \ldots, M_p\}$ setzen können.

■

Falls $\mathcal{O}$ Teilmenge von $\mathrm{Est}_\tau(D)$ für ein volles faktorielles Design D ist, können wir das lineare Modell zu $\mathcal{O}$ natürlich auf Grund von D bestimmen; falls $\mathcal{O}$ allerdings deutlich kleiner als $\mathrm{Est}_\tau(D)$ ist, sollte das auch mit deutlich geringerem Aufwand möglich sein, nämlich mit jeder Fraktion $\mathcal{F}$ von D mit $\mathrm{Est}_\tau(\mathcal{F}) \supseteq \mathcal{O}$. Es genügt, wenn wir die minimalen unter diesen Fraktionen bestimmen, denn jede andere enthält mindestens eine von diesen.

Via GRÖBNER-Basen und Monomordnungen lassen sich (wenn auch mit beträchtlichem Aufwand) alle diese Fraktionen bestimmen; wir wollen uns hier aber mit einem weniger ambitiösen Ziel begnügen und nur die Fraktionen bestimmen, die Nullstellenmengen von Idealen sind, die von Distraktionen von Polynomen erzeugt werden. Die Theorie dazu findet man bei

> LORENZO ROBBIANO, MARIA PIERA ROGANTIN: Full factorial designs and distracted fractions *in* BRUNO BUCHBERGER, FRANZ WINKLER [HRSG]: Gröbner bases and applications Linz, *Cambridge University Press,* 1998, Seite 473–482

wo auch Aussagen über die Anzahl der so zu findenden und der insgesamt existierenden Fraktionen $\mathcal{F}$ zu finden sind. Ansätze zur allgemeinen Lösung des Problems findet man in

> MASSIMO CARBOARA, LORENZO ROBBIANO: Families of Ideals in Statistics *in* KÜCHLIN [HRSG.]: Proceedings of the 1997 ISSAC, *ACM Press,* 1997, Seite 404–409

Wir gehen also aus von einem vollständigen faktoriellen Design D und wollen dort Teilmengen $\mathcal{F}$ bestimmen, die Nullstellenmengen von Distraktionen sind. Da wir D kennen, brauchen wir dazu nicht das volle Ideal $I(\mathcal{F})$, sondern es reicht, wenn wir Funktionen finden, die auf D genau in den Punkten von $\mathcal{F}$ verschwinden. Etwaige gemeinsame Nullstellen außerhalb von D interessieren uns nicht, und natürlich sind auch Polynome, die auf ganz D verschwinden, für uns uninteressant.

Zur Konstruktion verwenden wir Distraktionen der Monome aus Min($\mathcal{O}$), und die verschwinden genau dann sogar auf ganz D, wenn das zugehörige Monom bereit im Ideal $\mathrm{FM}_\tau\bigl(I(D)\bigr)$ der führenden Monome von $I(D)$ liegen. Da uns diese Monome nicht interessieren, entfernen wir sie aus Min($\mathcal{O}$):

Definition: CutOut($\mathcal{O}$) = Min($\mathcal{O}$) $\smallsetminus$ $\mathrm{FM}_\tau\bigl(I(D)\bigr)$

Damit gilt

Satz: $D = S_1 \times \cdots \times S_n$ sei ein volles faktorielles Design, G sei die universelle reduzierte GRÖBNER-Basis von $I(D)$, und $\mathcal{O}$ sei ein Ordnungsideal, das in $\mathrm{Est}_\tau(D)$ liegt mit CutOut($\mathcal{O}$) = $\{M_1, \dots, M_r\}$. Die Vektoren $a^{(i)}$ seien so definiert, daß ihre Komponenten gleich den verschiedenen Elementen von S_i in irgendeiner Reihenfolge sind. Dann bildet G zusammen mit den Polynomen $D(M_1), \dots, D(M_r)$ bezüglich jeder Monomordnung τ eine GRÖBNER-Basis eines Designideals einer Fraktion $\mathcal{F}$ von D mit $\mathrm{Est}_\tau(\mathcal{F}) = \mathcal{O}$.

Beweis: Wie wir bereits wissen, sind die Elemente der GRÖBNER-Basis G von $I(D)$ die Distraktionen der Monome $X_i^{\#S_i}$ bezüglich der $a^{(i)}$. Daher besteht die Vereinigung von G mit der Menge der M_j nur aus Distraktionen von Monomen bezüglich fester Vektoren und ist damit nach obigem Satz bezüglich jeder Monomordnung eine GRÖBNER-Basis des davon erzeugten Ideals I. Für $\mathcal{F} = V(I)$ wird $\mathrm{FM}_\tau(I)$ erzeugt von den führenden Monomen der Elemente von G und den M_i, also von CutOut($\mathcal{O}$) und $\mathrm{FM}_\tau\bigl(I(D)\bigr)$ und damit von $\mathrm{Min}_\tau(\mathcal{O})$. Somit ist $\mathrm{Est}_\tau(\mathcal{F}) = \mathcal{O}$.

■

Betrachten wir als Beispiel $D = \{0,1,2\} \times \{0,1\}^{10}$, das volle faktorielle Design für elf Faktoren, deren erster drei Stufen hat; alle weiteren haben nur zwei. D enthält somit $3 \times 2^{10} = 3\,072$ Punkte aus $\mathbb{R}^{11}$.

Wir erwarten nicht, daß jede der 3 072 Faktorkombinationen relevant ist und beschränken und auf das Modell

$$\mathcal{O} = \{1, X_1, \ldots, X_{11}, X_1X_2, X_1X_3, X_1X_4, X_1X_2X_3, X_2X_3, X_1^2\}\,.$$

Es gibt also nur eine Dreierinteraktion, und nur ein Monom kommt als Quadrat vor. Die bezüglich der Teilbarkeitsrelation maximalen Elemente von $\mathcal{O}$ sind $X_5, \ldots, X_{11}, X_1X_2X_3, X_1X_4$ und X_1^2. Die Menge Min($\mathcal{O}$) enthält somit alle Monome der Form MX_i, wobei M irgendeines dieser Monome ist.

$\mathrm{FM}_\tau\bigl(I(D)\bigr)$ wird erzeugt von X_1^3 und $X_2^2, \ldots, X_{10}^2$; CutOut($\mathcal{O}$) besteht also aus den davon verschiedenen Monomen aus Min($\mathcal{O}$).

Um kurze Polynome zu bekommen, wählen wir $a^{(1)} = (0,1,2)$ und $a^{(i)} = (0,1)$ für $i \geq 2$, d.h. wir stellen die Null an die erste Stelle. Die Distraktionen der Monome sind dann für alle Monome, die kein Quadrat enthalten, einfach die Monome selbst, und für die $X_1^2X_i$ sind es die Polynome $X_1(X_1-1)X_i$.

Der Nullpunkt und alle Punkte, bei denen genau eine Koordinate von Null verschieden ist, sind Nullstellen aller dieser Polynome; da x_1 auch den Wert zwei annehmen kann, sind dies schon einmal dreizehn Lösungen.

Falls ein x_i mit $i \geq 5$ von Null verschieden ist, müssen alle anderen x_j verschwinden, da X_iX_j eines der Erzeugenden des Ideals ist. Wegen der Polynome $D(X_1^2X_j) = X_1(X_1-1)X_j$ müssen auch im Falle $x_1 = 2$ alle x_j mit $j > 1$ verschwinden.

Ist $x_4 = 1$, so kann wegen der Polynome $X_1X_4X_j$ auch $x_1 = 1$ sein, falls alle anderen x_j verschwinden; es gibt also außer dem Einheitspunkt noch den Punkt $(1, 0, 0, 1, 0, \dots, 0)$.

Wenn alle x_i mit $i \geq 4$ verschwinden, ist für x_1 bis x_3 jede Kombination aus Nullen und Einsen möglich; dies ergibt acht Lösungen, von denen vier neu sind. Insgesamt haben wir also 18 Punkte, was erwartungsgemäß gleich der Elementanzahl von $\mathcal{O}$ ist.

§5: Berechnung der Est-Mengen

Nachdem wir uns im letzten Abschnitt damit beschäftigt haben, zu einem Ordnungsideal (und damit zu einem statistischen Modell) ein Design zu finden, mit dem sich die Modellparameter schätzen lassen, wollen wir nun zurückkehren zum Hauptthema dieser Vorlesung, also der Frage, welche Modelle wir auf der Grundlage eines vorgegebenen Design identifizieren können. Wir gehen also wieder aus von einem Design $D = \{x^{(1)}, \dots, x^{(r)}\} \subset k^n$ mit Ideal $I(D) \triangleleft k[X_1, \dots, X_n]$.

Beginnen wir mit eine festen Monomordnung τ auf $k[X_1, \dots, X_n]$. Mit den Methoden, die uns bislang zur Verfügung stehen, müssen wir zur Berechnung von $\mathrm{Est}_\tau(D)$ zunächst eine GRÖBNER-Basis von $I(D)$ bezüglich τ bestimmen und dann die Standardmonome dazu. Wir wollen uns überlegen, wie wir alternativ diese Standardmonome auch berechnen können, indem wir das Design sukzessive aufbauen. Dazu sei

$$D_m = \{x^{(1)}, \dots, x^{(m)}\} \quad \text{für} \quad m = 1, \dots, r\,,$$

und wir wollen nacheinander die Mengen $\mathrm{Est}_\tau(D_m)$ bestimmen.

Wie wir wissen, ist jede Est-Menge ein Ordnungsideal, enthält also die Eins, und außerdem hat $\mathrm{Est}_\tau(D_m)$ genauso viele Elemente wie D_m, also m Stück. Somit ist auf jeden Fall $D_1 = \{1\}$.

Lemma: Ist D ein Design und $D' \subset D$, so ist $\mathrm{Est}_\tau(D') \subset \mathrm{Est}_\tau(D)$.

Beweis: Für ein Design D besteht $\mathrm{Est}_\tau(D)$ aus allen Monomen, die durch kein führendes Monom eines Elements einer GRÖBNER-Basis teilbar sind. Da diese führenden Monome das Ideal der führenden Monome von $I(D)$ erzeugen, ist dies äquivalent dazu, daß die Monome aus $\mathrm{Est}_\tau(D)$ durch kein Monom aus $\mathrm{FM}_\tau\big(I(D)\big)$ teilbar sein dürfen.

Für eine Teilmenge $D' \subset D$ ist $I(D)$ eine Teilmenge von $I(D')$, also ist auch $\mathrm{FM}_\tau\big(I(D)\big)$ eine Teilmenge von $\mathrm{FM}_\tau\big(I(D')\big)$. Falls ein Monom durch keines der Monome aus $\mathrm{FM}_\tau\big(I(D')\big)$ teilbar ist, kann es also erst recht durch keines aus $\mathrm{FM}_\tau\big(I(D)\big)$ teilbar sein, was die Behauptung beweist.

■

Damit ist also $\mathrm{Est}_\tau(D_{m-1}) \subset \mathrm{Est}_\tau(D_m)$, und $\mathrm{Est}_\tau(D_m)$ hat genau ein Element mehr. Es gibt also ein Monom $M \notin \mathrm{Est}_\tau(D_{m-1})$, so daß $\mathrm{Est}_\tau(D_m) = \mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ ist. In §1 hatten wir zu einem Design D und einer Menge $\mathbb{F}$ von Funktionen f_i die Z-Matrix definiert als die Matrix mit Einträgen $z_{ij} = f_i\big(x^{(j)}\big)$. Da das lineare Modell mit den Monomen aus $\mathrm{Est}_\tau(D_m)$ als Basisfunktionen anhand von D_m eindeutig identifizierbar ist, muß die Z-Matrix zu D_m und $\mathrm{Est}_\tau(D_m)$ invertierbar sein. Außerdem ist $\mathrm{Est}_\tau(D_m)$ ein Ordnungsideal; daher muß M Produkt eines Monoms aus D_{m-1} mit einer der Variablen X_i sein.

Durch diese Bedingungen muß M noch nicht eindeutig bestimmt sein. Angenommen, es gibt zwei Monome M und M' derart, daß sowohl $\mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ als auch $\mathrm{Est}_\tau(D_{m-1}) \cup \{M'\}$ Ordnungsideale sind und für D_m auf invertierbare Z-Matrizen führen. Dann können wir die Monome M und M' bezüglich der Monomordnung τ miteinander vergleichen; M sei das kleinere der beiden Monome.

Wäre $\mathrm{Est}_\tau(D_m) = \mathrm{Est}_\tau(D_{m-1}) \cup \{M'\}$, so wäre M kein Standardmonom, wäre also teilbar durch das führende Monom eines Elements g der reduzierten GRÖBNER-Basis von $I(D_m)$ bezüglich τ. Da $\mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ ein Ordnungsideal ist, liegt jeder echte Teiler von M in $\mathrm{Est}_\tau(D_{m-1})$ und damit erst recht in $\mathrm{Est}_\tau(D_m)$. Somit muß M das führende Monom von g sein. Alle anderen Monome in g sind Standardmonome, liegen also in $\mathrm{Est}_\tau(D_{m-1}) \cup \{M'\}$. Das Monom M' kann allerdings nicht vorkommen, da das bezüglich τ kleinere Monom M führendes Monom von g ist. Somit ist g eine Linearkombination von Monomen aus dem Ordnungsideal $\mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$, und als Element von $I(D_m)$ verschwindet g natürlich auf ganz D_m. Wegen der Invertierbarkeit der Z-Matrix muß daher $g = 0$ sein, was für ein Element einer reduzierten GRÖBNER-Basis unmöglich ist. Also kann $\mathrm{Est}_\tau(D_m)$ nicht gleich $\mathrm{Est}_\tau(D_{m-1}) \cup \{M'\}$ sein.

Somit ist $\mathrm{Est}_\tau(D_m) = \mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ für das bezüglich τ kleinste Monom M, für das $\mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ ein Ordnungsideal ist und eine invertierbare Z-Matrix über D_m hat.

Dies führt auf folgenden Algorithmus zur Berechnung von $\mathrm{Est}_\tau(D)$:

1. Schritt: Setze $\mathrm{Est}_\tau(D_1) = \{1\}$

m-ter Schritt, $m \geq 2$: Setze $\mathrm{Est}_\tau(D_m) = \mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ für das bezüglich τ kleinste Monom M, für das $\mathrm{Est}_\tau(D_{m-1}) \cup \{M\}$ ein Ordnungsideal ist und eine invertierbare Z-Matrix über D_m hat.

Betrachten wir etwa $D = \big\{(0,0),\ (0,2),\ (1,1)\big\}$ für die lexikographische Ordnung mit $X > Y$ und die angegebene Reihenfolge der Punkte. Für $D_1 = \{(0,0)\}$ ist natürlich $\mathrm{Est}_\tau(D_1) = \{1\}$. Dies läßt sich auf zwei Arten zu einem zweielementigen Ordnungsideal erweitern: Entweder zu $\{1, X\}$ oder zu $\{1, Y\}$. Da beide Punkte von $D_2 = \big\{(0,0),\ (0,2)\big\}$ die gleiche x-Koordinate haben, führt $\{1, X\}$ zu einer singulären Z-Matrix; wir müssen also $\{1, Y\}$ nehmen, wofür wir die invertierbare Z-Matrix $\left(\begin{smallmatrix}1 & 1\\ 0 & 2\end{smallmatrix}\right)$ bekommen.

Dies wiederum läßt sich auch auf zwei Arten erweitern: Entweder zu $\{1, Y, Y^2\}$ oder zu $\{1, Y, X\}$. Wir erhalten bezüglich D die beiden nichtsingulären Z-Matrizen

$$\begin{pmatrix}1 & 0 & 0\\ 1 & 2 & 0\\ 1 & 0 & 1\end{pmatrix} \quad \text{und} \quad \begin{pmatrix}1 & 0 & 0\\ 1 & 2 & 4\\ 1 & 0 & 1\end{pmatrix}.$$

Da wir mit der lexikographischen Ordnung mit $X > Y$ arbeiten, ist $X > Y^2$, wir müssen also $\{1, Y, Y^2\}$ nehmen. Hätten wir stattdessen mit einer graduiert lexikographischen Ordnung gearbeitet, egal ob für $X > Y$ oder $Y > X$, wäre $Y^2 > X$, so daß wir uns dann für $\{1, Y, X\}$ entscheiden müßten. Dies illustriert, daß die lexikographische Ordnung Modelle bevorzugt, in denen eine (oder mehrere) „kleine“ Variablen dominieren, während die graduiert lexikographische Ordnung zu ausgewogeneren Modellen führt,wobei freilich auch hier die Reihenfolge der Variablen zusätzliche Prioritäten setzt. Andere Monomordnungen können zu Modellen mit wieder anderen Eigenschaften führen. Welche in einem konkreten Zusammenhang nützliche sind, hängt natürlich von der Anwendung ab: Manchmal dominiert eine der Eingabevariablen, in anderen Zusammenhängen sind alle ziemlich gleichberechtigt, und oft kommt es auch auf das Zusammenspiel verschiedener Variablen, d.h. auf gemischte Terme an. Die Tatsache, daß man anhand desselben Designs D ganz verschiedene Modelle identifizieren kann, bezeichnet man in der Statistik als *confounding* (Verwechseln). Welches der verschiedenen Modelle man in einem konkreten Zusammenhang nehmen sollte, ist keine mathematische Frage, sondern hängt ab vom eventuell vorhandenen Vorwissen über den zu untersuchenden Sachverhalt. Wenn wir zwei Modelle anhand von D und einer Funktion $D \to k$ identifiziert haben, muß die Differenz der beiden entstehenden Polynome natürlich im Designideal $I(D)$ liegen.

Genau wie man in der Numerik bei der Approximation einer transzendenten Funktion praktisch nie ein Interpolationspolynom betrachtet, das an vielen vorgegebenen Stellen exakt den korrekten Wert liefert, sucht man auch in der Statistik praktisch nie nach Modellen, die eine gemessene oder etwa durch Umfragen erhaltene Funktion $D \to k$ exakt darstellt: Stattdessen betrachtet man Modelle, die nur auf einer oft deutlich kleineren Teilmenge von $\mathrm{Est}_\tau(D)$ beruhen. Welche das sind, ist wieder teils abhängig von außermathematischem Vorwissen, teils hilft auch nur Experimentieren mit verschiedenen Ansätzen, für die man dann beispielsweise nach der Methode der kleinsten Quadrate entscheiden kann, welche davon gute Ergebnisse liefern.

Im Falle eines Designs mit minimalem Fächer ist das relativ einfach, da wir mit Teilmengen eines festen Ordnungsideals arbeiten können. Leider führt das aber auch nur auf wenige identifizierbare Modelle. Wenn wir kein Vorwissen über die Art des zu erwartenden Modells haben, ist es daher nützlicher, mit Designs zu arbeiten, deren Fächer möglichst groß ist.

Definition: *a)* Für zwei natürliche Zahlen r, n sei $\mathcal{G}(r, n)$ die Menge aller Ordnungsideale mit r Elementen in n Variablen.
b) Ein Design $D \subset k^n$ aus r Punkten heißt *Design mit maximalem Fächer,* wenn für jedes Ordnungsideal aus $\mathcal{G}(r, n)$ die Z-Matrix über D invertierbar ist.

$\mathcal{G}(r,n)$ ist natürlich stets eine endliche Menge, denn da ein Ordnungsideal mit jedem Element auch dessen sämtliche Teiler enthält, besteht jedes $\mathcal{O} \in \mathcal{G}(r,n)$ aus Monomen vom Grad kleiner r, und da die Menge aller Monome in n Variablen vom Grad kleiner r endlich ist, hat sie auch nur endlich viele Teilmengen.

Die Menge alle Designs aus r Punkten aus k^n ist $\bigl(k^n\bigr)^r$, was wir mit k^{nr} identifizieren können. Für jedes $\mathcal{O} \in \mathcal{G}(r,n)$ und jedes Design $D \in k^{rn}$ sind die Einträge der Z-Matrix Polynomfunktionen in den rn Koordinaten, und damit ist auch die Determinante der Z-Matrix eine Polynomfunktion auf k^{rn}. Die Designs mit einer singulären Z-Matrix bezüglich $\mathcal{O}$ liegen somit auf einer Hyperfläche (d.h. der Nullstellenmenge eines Polynoms) in k^{nr}, und die Menge aller Designs, die für *irgendein* $\mathcal{O} \in \mathcal{G}(r,n)$ zu eine singulären Z-Matrix führen, liegt in der Vereinigung endlich vieler solcher Hyperflächen.

Speziell im Fall $k = \mathbb{R}$ wissen wir, daß die Nullstellenmenge eines von Null verschiedenen Polynoms bezüglich des LEBESGUE-Maßes eine Nullmenge ist; die Designs aus $\mathbb{R}^{rn}$, die *keinen* maximalen Fächer haben, liegen also in einer Nullmenge. Wenn wir also zufällig ein Design aus $\mathbb{R}^{rn}$ oder (realistischer) einen beschränkten rn-dimensionalen Quader wählen, haben wir fast sicher ein Design mit maximalem Fächer. Leider hängt dieses Argument ganz entscheidend damit zusammen, daß wir über den rechnerisch eher unzugänglichen reellen Zahlen arbeiten; für uns interessante Teilmengen wie $\mathbb{Z}^{rn}$ und $\mathbb{Q}^{rn}$ sind allesamt Nullmengen, so daß wir dafür überhaupt keine Aussage bekommen. Soweit mir bekannt, gibt es nicht einmal eine Aussage darüber, ob es für jedes Paar (r,n) überhaupt ein Design $D \subset \mathbb{Z}^{rn}$ oder $D \subset \mathbb{Q}^{rn}$ mit maximalem Fächer gibt. Obwohl jedes „zufällig“ aus $\mathbb{R}^{rn}$ gewählte Design maximalen Fächer hat, gibt es meines Wissens auch keinen Algorithmus, der zu gegebenen Werten von r und n auch nur ein solches Design konstruiert.

Wie zu Beginn der Vorlesung bereits erwähnt, gibt es allerdings mathematische Methoden, oftmals algebraischer Art, zur Konstruktion von Designs, anhand derer man möglichst viel Information gewinnen kann. Diese zählen nicht zur algebraischen Statistik und sind vielfach auch deutlich älter als diese. Stattdessen bilden sie ein eigenes Teilgebiet der

Mathematik, das der optimalen Versuchsplanung. Solche Algorithmen können Designs mit nichtminimalem Fächer liefern, und wir möchten auch dazu die eindeutig identifizierbaren Modelle finden.

Glücklicherweise läßt sich der obige Algorithmus zur Bestimmung von $\mathrm{Est}_\tau(D)$ leicht umformulieren zu einem Algorithmus zur Bestimmung *aller* Ordnungsideale, die anhand von D identifiziert werden können.

Auch hier gehen wir wieder punktweise vor, und natürlich ist wieder $\mathrm{Est}_\tau(D_1) = \{1\}$. In den nächsten Schritten betrachten wir nun aber nicht nur das um ein Element größere Ordnungsideal mit dem bezüglich τ kleinstmöglichen neuen Monom, sondern wir betrachten *alle* Ordnungsideale, die durch Hinzufügen eines neuen Monoms entstehen und zu einer invertierbaren Z-Matrix führen. Für jedes der so entstehenden Ordnungsideale werden dann auch alle Folgeschritte durchgeführt. Dies kann zu einem exponentiell ansteigenden Aufwand führen, allerdings werden gelegentlich auch Ordnungsideale entstehen, die bereits vorher bekannt waren, da bei ihrer Konstruktion die gleichen Monome in verschiedener Reihenfolge auftraten. Falls wir diesen Algorithmus vollständig durchühren, erhalten wir am Ende eine Liste aller Ordnungsideale mit $\#D$ Elementen, die anhand von D identifiziert werden können.

Insbesondere können wir so auch entscheiden, ob ein Design maximalen Da sich Fächer hat: $\mathcal{G}(r,n)$ problemlos bestimmen wird, können wir leicht feststellen, ob alle seine Elemente in der durch den Algorithmus berechneten Liste vorkommen

§ 6: Nichtpolynomiale Modelle

Nicht alle Phänomene lassen sich polynomial modellieren: Bei periodischen Prozessen oder bei exponentiellem Wachstum etwa kann ein polynomialer Ansatz bestenfalls einen beschränkten Zeitraum approximativ beschreiben, und auch das nur mit großem Aufwand. Trotzdem läßt sich die hier behandelte Theorie manchmal auch in solchen Fällen anwenden.

Am einfachsten geht das im Falle von Modellen, deren Basisfunktionen geeignete Exponentialfunktionen sind: Wir betrachten für n feste, über $\mathbb{Q}$ linear unabhängige reelle Zahlen $\lambda_1, \dots, \lambda_n$ Funktionen, die Linearkombinationen von Termen der Form

$$e^{a_1\lambda_1 x_1 + \cdots + a_n\lambda_n x_n}$$

mit ganzen Zahlen a_i sind. Mit der Abkürzung Y_i für die Funktion

$$(x_1, \dots, x_n) \mapsto e^{\lambda_i x_i}$$

können wir so eine Funktion auch als Linearkombination von „Monomen" $Y_1^{a_1} \cdots Y_n^{a_n}$ schreiben, wobei allerdings im Gegensatz zu den „echten" Monomen auch negative Exponenten a_i zugelassen sind. Damit befinden wir uns fast im Polynomring $k[Y_1, \dots, Y_n]$, genauer gesagt in dessen Erweiterung

$$k[Y_1, \dots, Y_n, Y_1^{-1}, \dots, Y_n^{-1}]\,,$$

die wir vor allem deshalb brauchen, weil eine Exponentialfunktion und damit auch die Funktion Y_i den Wert Null nicht annehmen kann. Ansonsten können wir in diesem Ring weitgehend so rechnen, wie wir es von Polynomringen gewohnt sind und damit die in diesem Kapitel betrachtete Theorie auch auf solche Modelle erweitern.

Wenn wir den Körper k durch die imaginäre Einheit i erweitern zu $k(i) = k \oplus ki$, können wir auf diese Weise über die EULERschen Formeln

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin x = \frac{e^{-ix} - e^{-ix}}{2i}, \quad e^{ix} = cos x + i \sin x$$

auch Modelle für periodische Zusammenhänge behandeln.