Die Lineare Algebra-Methode

Mahir Kilic

23. Juni 2004

1 Einführung

1.1 Überblick

Im Allgemein benutzt man die Lineare Algebra-Methode in der Kombinatorik wie folgt:

Für die Bestimmung einer Obergrenze für die Mächtigkeit einer Menge werden die Elemente dieser Menge als Vektoren dargestellt. Dieser Vektorraum soll eine möglichst minimale Dimension haben. Zunächst zeigt man, dass diese Vektoren alle linear unabhängig sind. Hieraus kann man schlussfolgern, dass die Mächtigkeit der ursprünglichen Menge nicht die Dimension der Vektorraum überschreitet.

1.2 Lineare Algebra

Man benutzt ausschliesslich Lineare Algebra bei der Lineare Algebra-Methode. Deswegen ist einige Kentnisse der Lineare Algebra notwendig. Erst will ich ein Paar Definitionen geben, die für Verständnis der von mir gegebenen Beispielen notwendig sind.

1.2.1 Vektorräume, Untervektoräume, Lineare Abhängigkeit

Sei $\mathbf{F}$ ein Körper.

Definition 1 : Mit *Vektorraum* Ein Vektorraum V über $\mathbf{F}$ ist $\mathbf{F}^n$, wobei $v \in \mathbf{F}^n$ die Form $v = (v_1, \ldots, v_n)$ mit $v_i \in \mathbf{F}$ hat. In jedem solchen Vektorraum gibt es zwei grundlegenden Operationen: Komponentenweise Addition $u + v = (u_1 + v_1, \ldots, u_n + v_n)$ und Multiplikation mit einem Skalar $\lambda v = (\lambda v_1, \ldots, \lambda v_n), \lambda \in \mathbf{F}$.

Definition 2 : Ein Vektor v ist eine *Linearkombination* von den Vektoren $v_1, \ldots, v_m$, wenn der Vektor die Form: $v = \lambda_1 v_1 + \ldots + \lambda_m v_m$ mit $\lambda_i \in \mathbf{F}$ hat.

Definition 3 : *Ein Untervektorraum* von V ist eine nicht leere Untermenge von V, die unter Linearekombination abgeschlossen ist.

Definition 4 : Die *Lineare Hülle* ($span\{v_1, \ldots, v_m\}$) von $v_1, \ldots, v_m$ ist die Menge aller Linearkombinationen von $v_1, \ldots, v_m$.

Definition 5 : Ein Vektor u ist linear abhängig von den Vektoren $v_1, \ldots, v_m$, wenn $u \in span\{v_1, \ldots, v_m\}$. Andernfalls ist u von $v_1, \ldots, v_m$ unabhängig.

Definition 6 : $\lambda_1 v_1 + \ldots + \lambda_m v_m = 0$. Die Lösung von dieser Gleichung heisst trivial, wenn $\lambda_i = 0$ für alle i gilt.

$v_1, \ldots, v_m$ sind linear Unabhängig genau dann, wenn nur triviale Lösung der obigen Gleichung existiert.

Definition 7 : Eine *Basis* $B = \{v_1, \ldots, v_m\}$ von V ist eine Menge von unabhängigen Vektoren, welche V erzeugen. $V = span\{v_1, \ldots, v_m\}$

Definition 8 : Jede Basis von V hat die gleiche Kardinalität. Das ist die *Dimension* von V (dim(V)).

Behauptung 1. Wenn $(v_1, \ldots, v_k)$ linear unabhängige Vektoren in ein Vektorraum mit Dimension m sind, dann gilt $k \leq m$.

1.2.2 Skalarprodukt, Orthogonalität

Eine wichtige Operation in Vektoräume ist der Skalarprodukt.

Definition 9 : Gegeben sind zwei Vektoren $u = (u_1, \ldots, u_n)$ und $v = (v_1, \ldots, v_n)$. Ihre Skalarprodukt $\langle u, v \rangle$ ist so definiert:

$\langle u, v \rangle = u \cdot v := u_1 v_1 + \ldots + u_n v_n$.

Definition 10 : Vektoren u und v sind *orthogonal*, wenn $\langle u, v \rangle = 0$.

Definition 11 : Wenn $U \subseteq V$ ein Unterraum von V ist, dann ist das orthogonale Komplement wie folgt definiert:

$U^{\perp} = \{v \in V : \langle u, v \rangle = 0 \text{ für alle } u \in U\}$.

Behauptung 2. Sei V eine endlich dimensionale Vektorraum und $U \subseteq V$ ein Untervektorraum von V. Dann ist $dimU + dimU^{\perp} = dimV$.

1.2.3 Matrizen

Eine $m \times n - Matrix$ $\mathbf{A} = (a_{ij})$ über $\mathbf{F}$ ist ein rechteckiges Zahlenschema der Form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \ldots & a_{1n} \\ a_{21} & a_{22} \ldots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} \ldots & a_{mn} \end{pmatrix}$$

Wenn $A = \{a_{ij}\}$ ein $m \times n$ Matrix über Körper $\mathbf{F}$ und v ein Vektor in $\mathbf{F}^m$ ist, so ist vA ein Vektor in $\mathbf{F}^n$, dessen j'te Element Skalarprodukt von v mit der j'ten Spalte von A ist. Daher sind die Zeilen von A genau dann linear unabhängig, wenn $vA \neq 0$ für alle $v \neq 0$ ist.

Definition 12 : *Spaltenrang* von Matrix A ist die Dimension von dem Vektorraum, den Spalten von A aufspannen.*Zeilenrang* von Matrix A ist die Dimension von dem Vektorraum, den die Zeilen von A aufspannen. Ein Ergebnis der Matrixtheorie besagt: Zeilenrang und Spaltenrang von A sind gleich.Dieser Wert ist der *Rang* von A($rg(A)$ oder $rg_{\mathbf{F}}(A)$). Unten sind ein Paar Ungleichengen mit dem Rang:

$$rg(A) - rg(B) \leq rg(A+B) \leq rg(A) + rg(B) \tag{1}$$

$$rg(AB) \leq min\{rg(A), rg(B)\}. \tag{2}$$

1.2.4 Lineare Gleichungssysteme

Sei $x = (x_1, \ldots, x_n)$ ein unbekannter Vektor und $b = (b_1, \ldots, b_m) \in \mathbf{F}^m$ ein gegebener Vektor. So ist die Matrixgleichung $Ax = b$ ein Kurzform des Lineargleichungsystems.

$$a_{i1}x_1 + a_{i2}x_2 + \ldots, a_{in}x_n = b_i \quad \text{(i=1,...,m)}.$$

Für die Lösbarkeit von so ein Lineargleichungssystem haben wir eine nützliche Eigenschaft. Seien $a_1, \ldots, a_n \in \mathbf{F}^m$ Spalten von A. Dann ist $A \cdot x = x_1a_1 + \ldots + x_na_n$ Hiermit folgt, dass die Menge $\{A \cdot x : x \in \mathbf{F}^n\}$ in der Lineare Hülle der Spalten von A liegt. $A \cdot x = b$ ist genau dann lösbar, wenn $b \in span\{a_1, \ldots, a_n\}$ ist. $A \cdot x = b$ heisst homogen, wenn $b = 0$.

Behauptung 3. Sei A ein $m \times n$ Matrix über den Körper $\mathbf{F}$. Dann ist die Menge der Lösungen von LGS $A \cdot x = 0$ ein Untervektorraum von Vektorraum $\mathbf{F}^n$ mit der Dimension $n - rg(A)$.

1.2.5 Die Norm

Definition 13 : *Die Norm* eines Vektors v=$(v_1, \ldots, v_n)$ ist

$$\| v \| := \langle v, v \rangle^{1/2} = \left(\sum_{i=1}^n v_i^2\right)^{1/2}.$$

Die folgende Ungleichheit heisst Cauchy-Schwarz-Ungleichhung:
Behauptung 4.Für Vektoren $u, v \in \mathbf{R}^n$ gilt $\langle u, v \rangle \leq \| u \| \cdot \| v \|$.
Das heisst:

$$\left(\sum_{i=1}^n v_iu_i\right)^2 \leq \left(\sum_{i=1}^n v_i^2\right)\left(\sum_{i=1}^n u_i^2\right).$$

Beweis: Sei $\lambda \in \mathbf{R}$.

$0 \leq \langle \lambda u - v, \lambda u - v \rangle = \langle \lambda u, \lambda u - v \rangle - \langle v, \lambda u - v \rangle$

$= \lambda^2 \langle u, u \rangle - 2\lambda \langle u, v \rangle + \langle v, v \rangle$

Substitution $\lambda = \frac{\langle u,v \rangle}{\langle u,u \rangle}$

$0 \leq \frac{\langle u,v \rangle^2}{\langle u,u \rangle^2} \langle u, u \rangle - 2\frac{\langle u,v \rangle^2}{\langle u,u \rangle} + \langle v, v \rangle$

$= \langle v, v \rangle - \frac{\langle u,v \rangle^2}{\langle u,u \rangle}$

$\Rightarrow \langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle = \| u \| \| v \|.$

2 Ein Beispiel

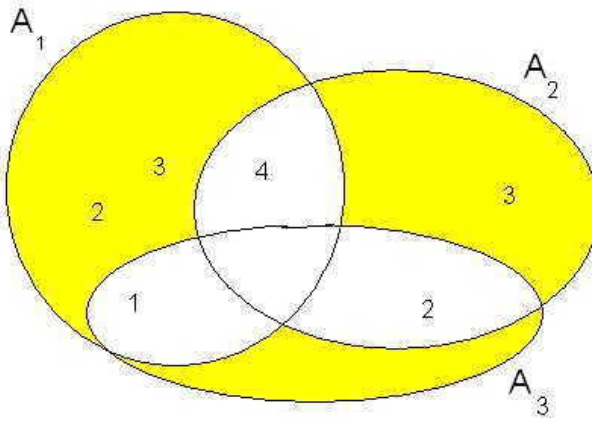

Abbildung 1: Ein Beispiel mit 3 Mengen und k=1

Fisher's Ungleichung Nehmen wir an, dass uns eine Familie von Mengen (F), die bestimmte Bedingungen erfüllen, gegeben ist. Wir möchten wissen, wie viele Mengen so eine Familie haben kann.

Satz. Seien $A_1, \ldots, A_m$ Untermengen von $\{1, \ldots, n\}$ und es gelte $|A_i \cap A_j| = k$ für eine k mit $1 \leq k \leq n$ für alle $i \neq j$.

Dann ist $m \leq n$.

(siehe Abbildung 1)

Dieses Theorem ist erst von dem Statistiker R.A.Fisher im Jahre 1940 für den Fall k=1 und alle Mengen A_i mit gleichen Mächtigkeit bewiesen. Im Jahre

1948 haben Brujin und Erdös den Satz für k=1 mit beliebiger Mächtigkeit der Mengen A_i bewiesen. Der ganz verallgemeinerte Fall ist erst im Jahre 1953 von Majumdar bewiesen worden. Der Beweis, den ich hier zeige, ist von Babai und Frankl.(1992)

Beweis. Seien $V_1, \ldots, V_m \in \{0,1\}^n$ Inzidenzvektoren von $A_1, \ldots, A_m$. Wegen Behauptung 1 ist es genug zu zeigen, dass diese Vektoren über $\mathbf{R}$ linear unabhängig sind. Nehmen wir das Gegenteil an. D.h die Gleichung $\sum_{i=1}^m \lambda_i v_i = 0$ eine nicht triviale Lösung hat. Offensichtlich ist $\langle v_i, v_i\rangle = |A_i|$ und $\langle v_i, v_j\rangle = k$, wenn $i \neq j$ ist. Daraus folgt:

$0 = (\sum_{i=1}^m \lambda_i v_i)(\sum_{j=1}^m \lambda_j v_j) = \sum_{i=1}^m \lambda_i^2 \langle v_i, v_i\rangle + \sum_{1\leq i\neq j\leq m} \lambda_i\lambda_j \langle v_i, v_j\rangle$
$= \sum_{i=1}^m \lambda_i^2 |A_i| + \sum_{1\leq i\neq j\leq m} \lambda_i\lambda_j k = \sum_{i=1}^m \lambda_i^2(|A_i| - k) + k\cdot(\sum_{i=1}^m \lambda_i)^2$

Klar ist $|A_i| \geq k$ für alle i. Hierbei gilt $|A_i| = k$ für maximal ein i. Sonst wäre die Schnittpunktbedingung nicht erfüllt. Hieraus folgt,dass rechte Seite grösser als 0 ist, was einen Widerspruch darstellt.

3 Flipping Card Games

Es gibt Situationen , wo allein Anwenden der lineare Algebra lassen sich interesante Ergebnisse erzielen. Vor allem können durch Linearkombinationen und/oder das Skalarprodukt dazu verwendet werden, um eine nützliche Information über die Inputvektoren zu kodieren. Das kann oft zu überraschenderweise effizienten Algorithmen führen .

Nehmen wir an, dass wir zwei Vektoren aus $\mathbf{F}_2^n$ $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ mit Länge n haben. Wir möchten entscheiden, ob $u = v$ ist. Aber unser Zugriff zu den Bits ist eingeschränkt.Zu jedem Zeitpunkt können wir zu jedem Index i entweder u_i oder v_i sehen. Wir können uns das so vorstellen, dass die Bits auf beiden Seiten von n Karten aufgetragen sind. Die Karten liegen auf dem Tisch. Wir können nur eine Seite jede Karte sehen.

(siehe Abbildung 2)

Ein Schritt besteht aus Drehung von einem oder mehreren Karten. Nach jedem Schritt können wir Informationen speichern, aber der Speicher ist nicht wiederverwendbar. Nach jedem Schritt müssen wir neuen Speicherplatz benutzen. Wir drehen die Karte, sehen die aktuelle Werte von Karten und benutzen die Informationen aus dem Speicher. Nachdem die Karte geschlossen wird, können wir eine Antwort geben oder schreiben zusätzliche Informationen in Speicher. Dann können wir mit dem nächsten Schritt anfangen.

Das Ziel ist es, mit möglichst wenig Speicherverbrauch zu entscheiden,

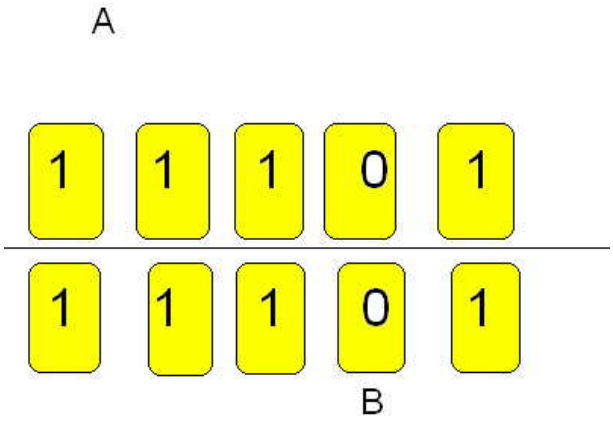

Abbildung 2: Beide Seiten von 5 Spielkarten

ob die beiden Seiten der Karten gleich sind. Speicherkapazität von n bits zu verbrauchen ist trivial für die Lösung. (Schreibe v einfach in den Speicher und drehe alle Karten um.) Kann man das besser machen?

Satz 6. Sei $n = r^2$ für eine $r \geq 1$. Man kann die Gleichheit von zwei Vektoren aus $\{0,1\}^n$ mit r+1 Schritten bestimmen, ohne mehr als r Bits zu schreiben.

Beweis. Teile die gegebene Vektoren u und v in r Teile der Länge r. $u = (u^1, \ldots, u^r)$ und $v = (v^1, \ldots, v^r)$ In dem ersten Schritt gucken wir den Vektor u an und rechnen den Vektor $w_0 := u^1 + u^2 + \ldots + u^r$ über $\mathbf{F_2}aus$. Diesen Vektor w_0 schreiben wir in den Speicher(Benutzung von r bits), und dann machen wir r Schritt wie folgt: Beim i'ten Versuch drehen wir nur die Karte, die im v^i sind und rechnen den Vektor

$w_i := u^1 + \ldots + u^{i-1} + v^i + u^{i+1} + \ldots + u^r$ aus.

Jetzt kontrollieren wir, ob der erhaltene Vektor w_i dem w_0 entspricht. Wenn alle Vektoren $w_1, \ldots, w_r$ mit dem w_0 zusammenfallen, sind u und v gleich. Sonst sind sie nicht gleich.Warum?

Wenn die Antwort $u = v$ ist, heisst das, dass nach der ersten Schritt gilt: $u^1+u^2+\ldots+u^r = v^1+u^2+\ldots+u^r$. Daraus folgt $u^1 = v^1$. Gleiches Argument gilt für jeden anderen Schritt. Daher ist $u = v$ korrekt.

Pudlak und Sgall(1997) haben gezeigt, dass eigentlich $O((\log n)^2)$ bits genug sind.

Satz 7. Es ist möglich, die Gleichheit von zwei Vektoren aus $\{0,1\}^n$ unter

Benutzung von nur $O(\log n)$ Schritten und $O(\log n)$ Speicherplatz pro Schritt zu testen.

Beweis. Jeder Schritt entspricht eine Untermenge I.($I \subseteq \{1, \ldots, n\}$) Nach jedem Schritt sehen wir n Bits: $|I|$ bits von u und $n - |I|$ bits von v. ($\{u_i : i \in I\}$ und $\{v_i : i \notin I\}$) Wir denken u und v als Vektoren in $\mathbf{R}^n$, die als Eintrag nur 0 und 1 haben. Idee ist (quadrat von) Euklidische Abstand von u und v zu rechnen und kontrollieren,ob es gleich 0 ist.

$$||u - v||^2 = \langle u, u\rangle + \langle v, v\rangle - 2\langle u, v\rangle$$

$$= \sum_{i=1}^n u_i^2 + \sum_{i=1}^n v_i^2 - 2\left((\sum_{i=1}^n u_i) \cdot (\sum_{i=1}^n v_i) - \sum_{i\neq j} u_i v_j\right)$$

Wir rechnen $\langle u, u\rangle$ und $\langle v, v\rangle$ je mit einem Schritt. ($I = \{1, \ldots, n\}, I = \oslash$) Wir brauchen $\log(n + 1) \uparrow$ bits Speicherplatz um diese zwei Zahlen, die zwischen 0 und n liegen, zu speichern. Es bleibt nur $\langle u, v\rangle = \sum_{i=1}^n u_i v_i$ zu rechnen.

Um das zu machen, rechnen wir erst die Produkt $N := (\sum_{i=1}^n u_i)(\sum_{i=1}^n v_i)$ mit Benutzung von gleichen Schritten und benutzen wir $2\log(n + 1) \uparrow$ bits Speicherplatz.(um den Produkt, die zwischen 0 und n^2 liegt, zu speichern) Um den gewünschte Produkt $\langle u, v\rangle$ zu rechnen, brauchen wir die Summe $\sum_{i\neq j} u_i v_j$. Das ist mit $2(\log(n) \uparrow)$ Schritten zu machen: Wähle die Proben, so dass für jede Kreuzterm eine andere berechnet werden kann und summiert alle Kreuzterme. Nach jedem Schritt schreib die Resultat mit $O(\log n)$ Speicherbits.