1 Optimierung, Lineare Algebra, Lineare und Ganzzahlige Programmierung : – Notation –

Prof. Dr. Martin Grötschel 3. Januar 2001

Der nachfolgende Text ist ein Auszug aus dem Skript zur Vorlesung "Lineare Optimierung", welche von mir im Wintersemester 2000/2001 gehalten wurde. Hier ist die Notation zur Optimierungstheorie zusammengefasst, die von mir verwendeten Bezeichnungen der Linearen Algebra sind ebenso aufgeführt, wie Bezeichnungsweisen für lineare und ganzzahlige Programme.

1.2 Optimierungsprobleme

Wir wollen zunächst ganz informell, ohne auf technische Spitzfindigkeiten einzugehen, Optimierungsprobleme einführen. Sehr viele Probleme lassen sich wie folgt formulieren.

Gegeben seien eine Menge S und eine geordnete Menge $(T, \leq)$, d. h. zwischen je zwei Elementen $s, t \in T$ gilt genau eine der folgenden Beziehungen $s < t$, $s > t$ oder $s = t$. Ferner sei eine Abbildung $f : S \to T$ gegeben. Gesucht ist ein Element $x^* \in S$ mit der Eigenschaft $f(x^*) \geq f(x)$ für alle $x \in S$ (Maximierungsproblem) oder $f(x^*) \leq f(x)$ für alle $x \in S$ (Minimierungsproblem). Es ist üblich hierfür eine der folgenden Schreibweisen zu benutzen:

$$(1.7) \qquad \begin{array}{lcl} \max_{x\in S} f(x) & \text{oder} & \max\{f(x) \mid x \in S\}, \\ \min_{x\in S} f(x) & \text{oder} & \min\{f(x) \mid x \in S\}. \end{array}$$

In der Praxis treten als geordnete Mengen $(T, \leq)$ meistens die reellen Zahlen $\mathbb{R}$, die rationalen Zahlen $\mathbb{Q}$ oder die ganzen Zahlen $\mathbb{Z}$ auf, alle mit der natürlichen Ordnung versehen (die wir deswegen auch gar nicht erst notieren). Die Aufgabe (1.7) ist viel zu allgemein, um darüber etwas Interessantes sagen zu können. Wenn S durch die Auflistung aller Elemente gegeben ist, ist das Problem entweder sinnlos oder trivial (man rechnet ganz einfach $f(x)$ für alle $x \in S$ aus). Das heißt,S muß irgendwie (explizit oder implizit) strukturiert sein, so daß vernünftige Aussagen über S möglich sind, ohne daß man alle Elemente in S einzeln kennt. Das gleiche gilt für die Funktion $f : S \to T$. Ist sie nur punktweise durch $x \mapsto f(x)$ gegeben, lohnt sich das Studium von (1.7) nicht. Erst wenn f durch hinreichend strukturierte "Formeln" bzw. "Eigenschaften" bestimmt ist, weden tieferliegende mathematische Einsichten möglich.

Die Optimierungsprobleme, die in der Praxis auftreten, haben fast alle irgendeine "vernünftige" Struktur. Das muß nicht unbedingt heißen, daß die Probleme dadurch auf einfache Weise lösbar sind, aber immerhin ist es meistens möglich, sie in das zur Zeit bekannte und untersuchte Universum der verschiedenen Typen von Optimierungsproblemen einzureihen und zu klassifizieren.

Im Laufe des Studiums werden Ihnen noch sehr unterschiedliche Optimierungsaufgaben begegnen. Viele werden von einem der folgenden Typen sein.

(1.8) Ein Kontrollproblem. Gegeben sei ein Steuerprozeß (z. B. die Bewegungsgleichung eines Autos), etwa der Form

$$\dot{x}(t) = f(t, x(t), u(t)),$$

wobei u eine Steuerung ist (Benzinzufuhr). Ferner seien eine Anfangsbedingung

$$x(0) = x_0,$$

(z. B.: das Auto steht) sowie eine Endbedingung

$$x(T) = x_1$$

(z.B. das Auto hat eine Geschwindigkeit von 50 km/h) gegeben. Gesucht ist eine Steuerung u für den Zeitraum $[0, T]$, so daß z. B.

$$\int_0^T |u|^2 \,\mathrm{dt}$$

minimal ist (etwa minimaler Benzinverbrauch).

□

(1.9) Ein Approximationsproblem. Gegeben sei eine (numerisch schwierig auszuwertende) Funktion f, finde eine Polynom p vom Grad n, so daß

$$||f - p|| \text{ oder } ||f - p||_\infty$$

minimal ist.

□

(1.10) Nichtlineares Optimierungsproblem. Es seien f, g_i $(i = 1, \ldots, m)$, h_j $(j = 1, \ldots, p)$ differenzierbare Funktionen von $\mathbb{R}^n \to \mathbb{R}$, dann heißt

$$\begin{array}{lll} \min f(x) & & \\ g_i(x) & \leq 0 & i = 1, \ldots, m \\ h_i(x) & = 0 & j = 1, \ldots, p \\ x & \in \mathbb{R}^n & \end{array}$$

ein nichtlineares Optimierungsproblem. Ist eine der Funktionen nicht differenzierbar, so spricht man von einem nichtdifferenzierbaren Optimierungsproblem. Im Allgemeinen wird davon ausgegangen, daß alle betrachteten Funktionen zumindest stetig sind.

□

(1.11) Konvexes Optimierungsproblem. Eine Menge $S \subseteq \mathbb{R}^n$ heißt **konvex**, falls gilt: Sind $x, y \in S$ und ist $\lambda \in \mathbb{R}$, $0 \leq \lambda \leq 1$, dann gilt $\lambda x + (1 - \lambda)y \in S$. Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt **konvex**, falls für alle $\lambda \in \mathbb{R}$, $0 \leq \lambda \leq 1$ und alle $x, y \in \mathbb{R}^n$ gilt

$$\lambda f(x) + (1 - \lambda)f(y) \geq f(\lambda x + (1 - \lambda)y).$$

Ist $S \subseteq \mathbb{R}^n$ konvex (z. B. kann S wie folgt gegeben sein $S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i = 1, \ldots, m\}$ wobei die g_i konvexe Funktionen sind), und ist $f : \mathbb{R}^n \to \mathbb{R}$ eine konvexe Funktion, dann ist

$$\min_{x \in S} f(x)$$

ein konvexes Minimierungsproblem.

□

(1.12) Lineares Optimierungsproblem (Lineares Programm).

Gegeben seien $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{(m,n)}$, $b \in \mathbb{R}^m$, dann heißt

$$\begin{aligned}&\max c^T x\\ &Ax \leq b\\ &x \in \mathbb{R}^n\end{aligned}$$

lineares Optimierungsproblem.

□

(1.13) Lineares ganzzahliges Optimierungsproblem.

Gegeben seien $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{(m,n)}$, $b \in \mathbb{R}^m$, dann heißt

$$\begin{aligned}&\max c^T x\\ &Ax \leq b\\ &x \in \mathbb{Z}^n\end{aligned}$$

lineares ganzzahliges (oder kurz: ganzzahliges) Optimierungsproblem.

□

Selbst bei Optimierungsproblemen wie $(1.8), \ldots, (1.13)$, die nicht sonderlich allgemein erscheinen mögen, kann es sein, daß (bei spezieller Wahl der Zielfunktion f und der Nebenbedingungen) die Aufgabenstellung (finde ein $x^* \in S$, so daß $f(x^*)$ so groß (oder klein) wie möglich ist) keine vernünftige Antwort besitzt. Es mag sein, daß f über S unbeschränkt ist; f kann beschränkt sein über S, aber ein Maximum kann innerhalb S nicht erreicht werden, d.h., das "max" müßte eigentlich durch "sup" ersetzt werden. S kann leer sein, ohne daß dies a priori klar ist, etc.. Der Leser möge sich Beispiele mit derartigen Eigenschaften überlegen! Bei der Formulierung von Problemen dieser Art muß man sich also Gedanken darüber machen, ob die betrachtete Fragestellung überhaupt eine sinnvolle Antwort erlaubt.

In unserer Vorlesung werden wir uns lediglich mit den Problemen (1.12) und (1.13) beschäftigen. Das lineare Optimierungsproblem (1.12) ist sicherlich das derzeit für die Praxis bedeutendste Problem, da sich außerordentlich viele und sehr unterschiedliche reale Probleme als lineare Programme formulieren lassen. Außerdem liegt eine sehr ausgefeilte Theorie vor, und mit den modernen Verfahren der linearen Optimierung können derartige Probleme mit Tausenden von Variablen und Ungleichungen gelöst werden. Dagegen ist Problem (1.13), viel schwieriger. Die

Beschränkung der Lösungsmenge auf die zulässigen ganzzahligen Lösungen führt direkt zu einem Sprung im Schwierigkeitsgrad des Problems. Verschiedene spezielle lineare ganzzahlige Programme können in beliebiger Größenordnung gelöst werden. Bei wenig strukturierten allgemeinen Problemen des Typs (1.13) versagen dagegen die Lösungsverfahren manchmal bereits bei weniger als 100 Variablen und Nebenbedingungen.

Über Kontrolltheorie (Probleme des Typs (1.8)), Approximationstheorie (Probleme des Typs (1.9)), Nichtlineare Optimierung (Probleme des Typs (1.10) und (1.11) werden an der TU Berlin Spezialvorlesungen angeboten. Es ist anzumerken, daß sich sowohl die Theorie als auch die Algorithmen zur Lösung von Problemen des Typs (1.8) bis (1.11) ganz erheblich von denen zur Lösung von Problemen des Typs (1.12) und (1.13) unterscheiden.

1.3 Notation

Wir gehen in der Vorlesung davon aus, daß die Hörer die Grundvorlesungen Lineare Algebra und Analysis kenen. Wir werden hauptsächlich Methoden der linearen Algebra benutzen und setzen eine gewisse Vertrautheit mit der Vektor- und Matrizenrechnung voraus.

Grundmengen

Wir benutzen folgende Bezeichnungen:

$\mathbb{N}$ $= \{1, 2, 3, \ldots\}$ $=$ Menge der **natürlichen Zahlen,**
$\mathbb{Z}$ $=$ Menge der **ganzen Zahlen,**
$\mathbb{Q}$ $=$ Menge der **rationalen Zahlen,**
$\mathbb{R}$ $=$ Menge der **reellen Zahlen,**

und mit M_+ bezeichnen wir die Menge der nichtnegativen Zahlen in M für $M \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. Wir betrachten $\mathbb{Q}$ und $\mathbb{R}$ als Körper mit der üblichen Addition und Multiplikation und der kanonischen Ordnung "$\leq$". Desgleichen betrachten wir $\mathbb{N}$ und $\mathbb{Z}$ als mit den üblichen Rechenarten versehen. Wir werden uns fast immer in diesen Zahlenuniversen bewegen, da diese die für die Praxis relevanten sind. Manche Sätze gelten jedoch nur, wenn wir uns auf $\mathbb{Q}$ oder $\mathbb{R}$ beschränken. Um hier eine saubere Trennung zu haben, treffen wir die folgende Konvention. Wenn wir das Symbol

$$\mathbb{K}$$

benutzen, so heißt dies immer das $\mathbb{K}$ einer der angeordneten Körper $\mathbb{R}$ oder $\mathbb{Q}$ ist. Sollte ein Satz nur für $\mathbb{R}$ oder nur für $\mathbb{Q}$ gelten, so treffen wir die jeweils notwendige Einschränkung.

Für diejenigen, die sich für möglichst allgemeine Sätze interessieren, sei an dieser Stelle folgendes vermerkt. Jeder der nachfolgend angegebenen Sätze bleibt ein wahrer Satz, wenn wir als Grundkörper $\mathbb{K}$ einen archimedisch angeordneten Körper wählen. Ein bekannter Satz besagt, daß jeder archimedisch angeordnete Körper isomorph zu einem Unterkörper von $\mathbb{R}$ ist, der $\mathbb{Q}$

enthält. Unsere Sätze bleiben also richtig, wenn wir statt $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}\}$ irgendeinen archimedisch angeordneten Körper $\mathbb{K}$ mit $\mathbb{Q} \subseteq \mathbb{K} \subseteq \mathbb{R}$ wählen.

Wir können in fast allen Sätzen (insbesondere bei denen, die keine Ganzzahligkeitsbedingungen haben) auch die Voraussetzung "archimedisch" fallen lassen, d. h. fast alle Sätze gelten auch für angeordnete Körper. Vieles was wir im $\mathbb{K}^n$ beweisen, ist auch in beliebigen metrischen Räumen oder Räumen mit anderen als euklidischen Skalarprodukten richtig. Diejenigen, die Spaß an derartigen Verallgemeinerungen haben, sind eingeladen, die entsprechenden Beweise in die allgemeinere Sprache zu übertragen.

In dieser Vorlesung interessieren wir uns für so allgemeine Strukturen nicht. Wir verbleiben in den (für die Praxis besonders wichtigen) Räumen, die über den reellen oder rationalen Zahlen errichtet werden. Also, nochmals, wenn immer wir das Symbol $\mathbb{K}$ im weiteren gebrauchen, gilt

$$\mathbb{K} \in \{\mathbb{R}, \mathbb{Q}\},$$

und $\mathbb{K}$ ist ein Körper mit den üblichen Rechenoperationen und Strukturen. Natürlich ist $\mathbb{K}_+ = \{x \in \mathbb{K} \mid x \geq 0\}$

Die Teilmengenbeziehung zwischen zwei Mengen M und N bezeichnen wir wie üblich mit $M \subseteq N$. Gilt $M \subseteq N$ und $M \neq N$, so schreiben wir $M \subset N$. $M \setminus N$ bezeichnet die mengentheoretische Differenz $\{x \in M \mid x \notin N\}$.

Vektoren und Matrizen

Ist R eine beliebige Menge, $n \in \mathbb{N}$, so bezeichnen wir mit

$$R^n$$

die Menge aller **n-Tupel** oder Vektoren der Länge n mit Komponenten aus R. (Aus technischen Gründen ist es gelegentlich nützlich, Vektoren $x \in R^0$, also Vektoren ohne Komponenten, zu benutzen. Wenn wir dies tun, werden wir es explizit erwähnen, andernfalls setzen wir immer $n \geq 1$ voraus.) Wir betrachten Vektoren $x = (x_i)_{i=1,\ldots,n} \in R^n$ immer als **Spaltenvektoren** d. h.

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

wollen wir mit Zeilenvektoren rechnen, so schreiben wir x^T (lies: x **transponiert**). Die Menge $\mathbb{K}^n$ ist bekanntlich ein n-dimensionaler Vektorraum über $\mathbb{K}$. Mit

$$y^T x := \sum_{i=1}^{n} x_i y_i$$

bezeichnen wir das **innere Produkt** zweier Vektoren $x, y \in \mathbb{K}^n$. Wir nennen x und y **senkrecht (orthogonal)**, falls $x^T y = 0$ gilt. Der $\mathbb{K}^n$ ist für uns immer (wenn nichts anderes gesagt wird) mit der **euklidischen Norm**

$$\|x\| := \sqrt{x^T x}$$

ausgestattet.

Für Mengen $S, T \subseteq \mathbb{K}^n$ und $\alpha \in \mathbb{K}$ benutzen wir die folgenden Standardbezeichnungen für Mengenoperationen

$$\begin{array}{ll} S+T & := \{x+y \in \mathbb{K}^n \mid x \in S, y \in T\}, \\ S-T & := \{x-y \in \mathbb{K}^n \mid x \in S, y \in T\}, \\ \alpha S & := \{\alpha x \in \mathbb{K}^n \mid x \in S, \}. \end{array}$$

Einige Vektoren aus $\mathbb{K}^n$ werden häufig auftreten, weswegen wir sie mit besonderen Symbolen bezeichnen. Mit e_j bezeichnen wir den Vektor aus $\mathbb{K}^n$, dessen j-te Komponente 1 und dessen übrige Komponenten 0 sind. Mit 0 bezeichnen wir den Nullvektor, mit **1** den Vektor, dessen Komponenten alle 1 sind. Also

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0\,\mathbf{1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ \mathbf{1} \end{pmatrix}.$$

Welche Dimension die Vektoren e_j, 0, **1** haben, ergibt sich jeweils aus dem Zusammenhang.

Für eine Menge R und $m, n \in \mathbb{N}$ bezeichnet

$$R^{(m,n)} \text{ oder } R^{(m \times n)}$$

die Menge der (m,n)-**Matrizen** (m Zeilen, n Spalten) mit Einträgen aus R. (Aus technischen Gründen werden wir gelegentlich auch $n = 0$ oder $m = 0$ zulassen, d. h. wir werden auch Matrizen mit m Zeilen und ohne Spalten bzw. n Spalten und ohne Zeilen betrachten. Dieser Fall wird jedoch immer explizit erwähnt, somit ist immer $n \geq 1$ und $m \geq 1$ vorausgesetzt.) Ist $A \in R^{(m,n)}$, so schreiben wir

$$A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

und meinen damit, daß A die folgende Form hat

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

Wenn nichts anderes gesagt wird, hat A die Zeilenindexmenge $M = \{1, \ldots, m\}$ und die Spaltenindexmenge $N = \{1, \ldots, n\}$. Die j-**te Spalte** von A ist ein m-Vektor, den wir mit $A_{.j}$ bezeichnen,

$$A_{.j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Die i**-te Zeile** von A ist ein Zeilenvektor der Länge n, den wir mit $A_{i.}$ bezeichnen, d. h.

$$A_{i.} = (a_{i1}, a_{i2}, \ldots, a_{in}).$$

Wir werden in dieser Vorlesung sehr häufig Untermatrizen von Matrizen konstruieren, sie umsortieren und mit ihnen rechnen müssen. Um dies ohne Zweideutigkeiten durchführen zu können, führen wir zusätzlich eine etwas exaktere als die oben eingeführte Standardbezeichnungsweise ein.

Wir werden — wenn immer es aus technischen Gründen notwendig erscheint — die Zeilen- und Spaltenindexmengen einer (m, n)-Matrix A nicht als Mengen sondern als Vektoren auffassen. Wir sprechen dann vom

$$\begin{array}{ll} \text{vollen Zeilenindexvektor} & M = (1, 2, \ldots, m) \\ \text{vollen Spaltenindexvektor} & N = (1, 2, \ldots, n) \end{array}$$

von A. Ein **Zeilenindexvektor** von A ist ein Vektor mit höchstens m Komponenten, der aus M durch Weglassen einiger Komponenten von M und Permutation der übrigen Komponenten entsteht. Analog entsteht ein **Spaltenindexvektor** durch Weglassen von Komponenten von N und Permutation der übrigen Komponenten. Ist also

$$\begin{array}{ll} I = (i_1, i_2, \ldots, i_p) & \text{ein Zeilenindexvektor von } A \text{ und} \\ J = (j_1, j_2, \ldots, j_q) & \text{ein Spaltenindexvektor von } A, \end{array}$$

so gilt immer $i_s, i_t \in \{1, \ldots, m\}$ und $i_s \neq i_t$ für $1 \leq s < t \leq m$, und analog gilt $j_s, j_t \in \{1, \ldots, n\}$ und $j_s \neq j_t$ für $1 \leq s < t \leq n$. Wir setzen

$$A_{IJ} := \begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_q} \\ a_{i_2 j_1} & a_{i_2 j_2} & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p j_1} & a_{i_p j_2} & \ldots & a_{i_p j_q} \end{pmatrix}$$

und nennen A_{IJ} **Untermatrix** von A.

A_{IJ} ist also eine (p, q)-Matrix, die aus A dadurch entsteht, daß man die Zeilen, die zu Indizes gehören, die nicht in I enthalten sind, und die Spalten, die zu Indizes gehören, die nicht in J enthalten sind, streicht und dann die so entstehende Matrix umsortiert.

Ist $I = (i)$ und $J = (j)$, so erhalten wir zwei Darstellungsweisen für Zeilen bzw. Spalten von A:

$$\begin{aligned} A_{IN} &= A_{i.}\ , \\ A_{MJ} &= A_{.j}\ . \end{aligned}$$

Aus Gründen der Notationsvereinfachung werden wir auch die folgende (etwas unsaubere) Schreibweise benutzen. Ist M der volle Zeilenindexvektor von A und I ein Zeilenindexvektor, so schreiben wir auch

$$I \subseteq M,$$

obwohl I und M keine Mengen sind, und für $i \in \{1, \ldots, m\}$ benutzen wir

$$i \in I \quad \text{oder } i \notin I,$$

um festzustellen, daß i als Komponente von I auftritt oder nicht. Analog verfahren wir bezüglich der Spaltenindizes.

Gelegentlich spielt die tatsächliche Anordnung der Zeilen und Spalten keine Rolle. Wenn also z. B. $I \subseteq \{1, \ldots, n\}$ und $J \subseteq \{1, \ldots, m\}$ gilt, dann werden wir auch einfach schreiben

$$A_{IJ},$$

obwohl diese Matrix dann nur bis auf Zeilen- und Spaltenpermutationen definiert ist. Wir werden versuchen, diese Bezeichnungen immer so zu verwenden, daß keine Zweideutigkeiten auftreten. Deshalb treffen wir ab jetzt die Verabredung, daß wir — falls wir Mengen (und nicht Indexvektoren) I und J benutzen – die Elemente von $I = \{i_1, \ldots, i_p\}$ und von $J = \{j_1, \ldots, j_q\}$ kanonisch anordnen, d. h. die Indizierung der Elemente von I und J sei so gewählt, daß $i_1 < i_2 < \ldots < i_p$ und $j_1 < j_2 < \ldots < j_q$ gilt. Bei dieser Verabredung ist dann A_{IJ} die Matrix die aus A durch Streichen der Zeilen i, $i \notin I$, und der Spalten j, $j \notin J$, entsteht.

Ist $I \subseteq \{1, \ldots, m\}$ und $J \subseteq \{1, \ldots, n\}$ oder sind I und J Zeilen- bzw. Splaltenindexvektoren, dann schreiben wir auch

$$A_{I.} \quad \text{statt} \quad A_{IN}$$

$$A_{.J} \quad \text{statt} \quad A_{MJ}$$

$A_{I.}$ entsteht also aus A durch Streichen der Zeilen i, $i \notin I$, $A_{.J}$ durch Streichen der Spalten j, $j \notin J$.

Sind $A, B \in \mathbb{K}^{(m,n)}$, $C \in \mathbb{K}^{(n,s)}$, $\alpha \in \mathbb{K}$, so sind

$$\begin{array}{rl} \text{die } \textbf{Summe} & A + B\ , \\ \text{das } \textbf{Produkt} & \alpha A, \\ \text{das } \textbf{Matrixprodukt} & AC \end{array}$$

wie in der linearen Algebra üblich definiert.

Für einige häufig auftretende Matrizen haben wir spezielle Symbole reserviert. Mit 0 bezeichnen wir die Nullmatrix (alle Matrixelemente sind Null), wobei sich die Dimension der Nullmatrix jeweils aus dem Zusammenhang ergibt. (Das Symbol 0 kann also sowohl eine Zahl, einen Vektor als auch eine Matrix bezeichnen). Mit I bezeichnen wir die Einheitsmatrix. Diese Matrix ist quadratisch, die Hauptdiagonalelemente von I sind Eins, alle übrigen Null. Wollen wir die Dimension von I betonen, so schreiben wir auch I_n und meinen damit die (n, n)-Einheitsmatrix. Diejenige (m, n)-Matrix, bei der alle Elemente Eins sind, bezeichnen wir mit E. Wir schreiben auch $E_{m,n}$ bzw. E_n, um die Dimension zu spezifizieren (E_n ist eine (n, n)-Matrix). Ist x ein

n-Vektor, so bezeichnet $\operatorname{diag}(x)$ diejenige (n,n)-Matrix $A = (a_{ij})$ mit $a_{ii} = x_i$ $(i = 1, \ldots, n)$ und $a_{ij} = 0$ $(i \neq j)$. Wir halten an dieser Stelle noch einmal folgendes fest. Wenn wir von einer Matrix A sprechen, ohne anzugeben, welche Dimension sie hat und aus welchem Bereich sie ist, dann nehmen wir implizit an, daß gilt $A \in \mathbb{K}^{(m,n)}$. Analog gilt immer $x \in \mathbb{K}^n$, wenn sich nicht aus dem Zusammenhang anderes ergibt.

Kombinationen von Vektoren, Hüllen, Unabhängigkeit

Ein Vektor $x \in \mathbb{K}^n$ heißt **Linearkombination** der Vektoren $x_1, \ldots, x_k \in \mathbb{K}^n$, falls es einen Vektor $\lambda = (\lambda_1, \ldots, \lambda_k)^T \in \mathbb{K}^k$ gibt mit

$$x = \sum_{i=1}^{k} \lambda_i x_i \ .$$

Gilt zusätzlich

$$\left.\begin{array}{l} \lambda \geq 0 \\ \lambda^T \mathbf{1} = 1 \\ \lambda \geq 0 \text{ und } \lambda^T \mathbf{1} = 1 \end{array}\right\} \quad \text{so heißt} \quad x \quad \left\{\begin{array}{l} \textbf{konische} \\ \textbf{affine} \\ \textbf{konvexe} \end{array}\right\} \quad \textbf{Kombination}$$

der Vektoren $x_1, \ldots, x_k$. Diese Kombinationen heißen **echt**, falls weder $\lambda = 0$ noch $\lambda = e_j$ für ein $j \in \{1, \ldots, k\}$ gilt.

Für eine nichtleere Teilmenge $S \subseteq \mathbb{K}^n$ heißt

$$\left.\begin{array}{l} \operatorname{lin}(S) \\ \operatorname{cone}(S) \\ \operatorname{aff}(S) \\ \operatorname{conv}(S) \end{array}\right\} \quad \text{die} \quad \left\{\begin{array}{l} \textbf{lineare} \\ \textbf{konische} \\ \textbf{affine} \\ \textbf{konvexe} \end{array}\right\} \quad \textbf{Hülle} \text{ von S, d. h.}$$

die Menge aller Vektoren, die als lineare (konische, affine oder konvexe) Kombination von endlich vielen Vektoren aus S dargestellt werden können. Wir setzen außerdem

$$\begin{array}{ll} \operatorname{lin}(\emptyset) & := \operatorname{cone}(\emptyset) := \{0\}, \\ \operatorname{aff}(\emptyset) & := \operatorname{conv}(\emptyset) := \emptyset. \end{array}$$

Ist A eine (m,n)-Matrix, so schreiben wir auch

$$\operatorname{lin}(A),\ \operatorname{cone}(A),\ \operatorname{aff}(A),\ \operatorname{conv}(A)$$

und meinen damit die lineare, konische, affine bzw. konvexe Hülle der Spaltenvektoren $A_{.1}, A_{.2}, \ldots, A_{.n}$ von A. Eine Teilmenge $S \subseteq \mathbb{K}^n$ heißt

$$\left\{\begin{array}{l} \textbf{linearer Raum} \\ \textbf{Kegel} \\ \textbf{affiner Raum} \\ \textbf{konvexe Menge} \end{array}\right\} \quad \text{falls} \quad \left\{\begin{array}{l} S = \operatorname{lin}(S) \\ S = \operatorname{cone}(S) \\ S = \operatorname{aff}(S) \\ S = \operatorname{conv}(S) \end{array}\right\}.$$

Eine nichtleere endliche Teilmenge $S \subseteq \mathbb{K}^n$ heißt **linear** (bzw. **affin**) **unabhängig**, falls kein Element von S als echte Linearkombination (bzw. Affinkombination) von Elementen von S

dargestellt werden kann. Die leere Menge ist affin, jedoch nicht linear unabhängig. Jede Menge $S \subseteq \mathbb{K}^n$, die nicht linear bzw. affin unabhängig ist, heißt **linear** bzw. **affin abhängig**. Aus der linearen Algebra wissen wir, daß eine linear (bzw. affin) unabhängige Teilmenge des $\mathbb{K}^n$ höchstens n (bzw. $n + 1$) Elemente enthält. Für eine Teilmenge $S \subseteq \mathbb{K}^n$ heißt die Kardinalität einer größten linear (bzw. affin) unabhängigen Teilmenge von S der **Rang** (bzw. **affine Rang**) von S. Wir schreiben dafür rang(S) bzw. arang(S). Die **Dimension** einer Teilmenge $S \subseteq \mathbb{K}^n$, Bezeichnung: $\dim(S)$, ist die Kardinalität einer größten affin unabhängigen Teilmenge von S minus 1, d. h. $\dim(S) =$ arang$(S) - 1$.

Der **Rang einer Matrix** A, bezeichnet mit rang(A), ist der Rang ihrer Spaltenvektoren. Aus der linearen Algebra wissen wir, daß rang(A) mit dem Rang der Zeilenvektoren von A übereinstimmt. Gilt für eine (m, n)-Matrix A, rang$(A) = \min\{m, n\}$, so sagen wir, daß A **vollen Rang** hat. Eine (n, n)-Matrix mit vollem Rang ist **regulär**, d. h. sie besitzt eine (eindeutig bestimmte) inverse Matrix (geschrieben A^{-1}) mit der Eigenschaft $AA^{-1} = I$.