Definition B.57 (Transposition)

Eine einfache aber wichtige Operation auf Matrizen ist die **Transposition**, die aus einer $(m \times n)$ Matrix A eine $(n \times m)$ Matrix $B = A^T$ macht. Hierbei gilt $\beta_{i\,j} = \alpha_{j\,i}$, so daß in Matrixschreibweise

$$A^T \;=\; \begin{pmatrix} \alpha_{1\,1} & \alpha_{2\,1} & \cdots & \alpha_{m\,1} \\ \alpha_{1\,2} & \alpha_{2\,2} & \cdots & \alpha_{m\,2} \\ & \cdots\cdots\cdots & & \\ \alpha_{1\,n} & \alpha_{2\,n} & \cdots & \alpha_{m\,n} \end{pmatrix} \;=\; (\beta_{i\,j})_{i=1\ldots n}^{j=1\ldots m} \,.$$

Bemerkung:

Nur die Diagonalelemente $(\alpha_{i\,i})_{i=1\ldots\,\min(m,n)}$ bleiben bei der Transposition unverändert, die anderen Elemente tauschen den Platz mit ihrem Gegenüber auf der anderen Seite der Diagonalen.

Lemma B.58 (Transpositionsregeln)

Man kann sich leicht davon überzeugen, daß die folgenden Regeln für das Transponieren gelten:

$$\begin{aligned} (A^T)^T &= A \\ (A+B)^T &= A^T + B^T \\ (\lambda A)^T &= \lambda A^T \\ (A\,B)^T &= B^T A^T \quad . \end{aligned}$$

Bemerkung:

Die Transposition ist also eine lineare Abbildung von $\mathbb{R}^{m\times n}$ nach $\mathbb{R}^{n\times m}$ und als solche sogar ihre eigene Inverse. Die letzte Gleichung bedeutet, daß die Transponierte eines Produktes gleich dem Produkt der transponierten Faktoren in umgekehrter Reihenfolge ist. Hierbei müssen wir natürlich wieder davon ausgehen, daß die Formate der Faktoren bezüglich der Produktbildung verträglich sind, was dann entsprechend für die Transponierten folgt.

Spezielle Matrixformen

Je nach ihrem Format, der Verteilung nicht verschwindender Elemente und gewissen algebraischen Eigenschaften unterscheidet man die folgenden häufig auftretenden Matrix Typen.

Zeilenvektor

$$A \in \mathbb{R}^{1\times n} \Rightarrow A = (\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1n})$$

In diesem Falle nennt man A einen **Zeilenvektor**.

Spaltenvektor

$$A \in \mathbb{R}^{m\times 1} \Rightarrow A = \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}$$

In diesem Falle nennt man A einen **Spaltenvektor**. Er kann von links mit einer m-spaltigen Matrix multipliziert werden, in diesem Fall stimmt das Matrix–Vektor–Produkt und das übliche Matrix–Matrix–Produkt überein.

Äusseres oder dyadisches Produkt

Das Produkt eines Zeilenvektors $\mathbf{a}^T = [(\alpha_i)_{i=1\ldots n}]^T \in \mathbb{R}^{1\times n}$ mit einem Spaltenvektor $\mathbf{b} = (\beta_i)_{i=1\ldots m} \in \mathbb{R}^{m\times 1}$ der gleichen Länge $m = n$ ergibt

$$\mathbf{a}^T\mathbf{b} = (\mathbf{a} * \mathbf{b}) = \mathbf{b}^T\mathbf{a} = \sum_{i=1}^{n} \alpha_i \beta_i \in \mathbb{R}^{1\times 1} \,.$$

Diese 1×1 Matrix kann man also als Skalar mit dem inneren Produkt zwischen $\mathbf{a}$ und $\mathbf{b}$ identifizieren. Wechselt man jedoch die Reihenfolge der Faktoren, so ergibt sich auch fuer $n \neq m$ die wohldefinierte Matrix

$$\mathbf{b}\mathbf{a}^T = (b_i a_j)_{j=1\ldots n}^{i=1\ldots m} \in \mathbb{R}^{m\times n}$$

Diese nennt man auch das **äussere** oder **dyadische Produkt** von **a** und **b**.

Verbilligte Produkte

Normalerweise kostet für $A \in \mathbb{R}^{m\times n}$ die Berechnung des Produktes $A\mathbf{v}$ mit einem Vektor $\mathbf{v} \in R^n$ genau $m \cdot n$ skalare Multiplikationen. Ist jedoch $A = \mathbf{b}\mathbf{a}^T$ ein äusseres Produkt so berechnet man viel billiger

$$A\mathbf{v} \;=\; (\mathbf{b}\mathbf{a}^T)\mathbf{v} \;=\; \mathbf{b}(\mathbf{a}^T v)\,.$$

Beachte, dass $\mathbf{b}(\mathbf{a}^T\mathbf{v})$ durch Bildung des Inneren Produktes $\mathbf{a}^T\mathbf{v} = \mathbf{a}\cdot\mathbf{v}$ und seine anschliessende Multiplikation mit $\mathbf{b}$ nur $n+m$ skalare Multiplikationen verlangt. Demgegenüber kostet alleine die explizite Berechnung des äusseren Produktes $\mathbf{b}\mathbf{a}^T$ genau $m \cdot n$ Multiplikationen. Entsprechend berechnet man das Produkt mit einer Matrix $V \in \mathbb{R}^{n\times p}$ als

$$(\mathbf{b}\mathbf{a}^T)V \;=\; \mathbf{b}(\mathbf{a}^T V) \;= \mathbf{b}(V^T\mathbf{a})^T$$

Die Produktbildung $\mathbf{b}(V^T\mathbf{a})^T$ kostet nur $(m+n)\cdot p$ skalare Multiplikationen während die Berechnung in der Form $(\mathbf{b}\mathbf{a}^T)V$ mehr als $m \cdot n \cdot p$ solche Operationen verlangt. Allgemeiner bezeichnet man die Fragestellung, in welcher Reihenfolge ein Produkt mehrerer Matrizen am billigsten berechnet werden kann, als **Matrixketten-Problem**. Es kann

Quadratische Matrix

$$A \in \mathbb{R}^{n\times n} \Rightarrow A^T \in \mathbb{R}^{n\times n}$$

Eine Matrix, deren Zeilenzahl gleich ihrer Spaltenzahl ist, heißt **quadratisch**. Alle linearen Abbildungen eines Raumes in sich selbst werden durch quadratische Matrizen beschrieben.

Symmetrische Matrix

$$A^T \;=\; A \in \mathbb{R}^{n\times n}$$

Quadratische Matrizen, die bezüglich der Transposition invariant sind, heißen **symmetrisch**. Diese bilden einen Unterraum von $\mathbb{R}^{n\times n}$. Dieser Unterraum hat die Dimension $n(n+1)/2$, da man lediglich die n Elemente in der Diagonale und entweder die $n(n-1)/2$ Elemente darüber oder die gleiche Zahl darunter frei wählen kann.

Schief symmetrische Matrix

$$A^T = -A \in \mathbb{R}^{n\times n}$$

Quadratische Matrizen mit dieser Eigenschaft heißen **schief symmetrisch**. Wie wir später sehen werden, sind alle ihre Eigenwerte rein imaginär.

Für jede quadratische Matrix gilt

$$A \;=\; \underbrace{\tfrac{1}{2}(A+A^T)}_{symmetrisch} + \underbrace{\tfrac{1}{2}(A-A^T)}_{schiefsymmetrisch} \quad .$$

Diese additive Zerlegung ist allerdings nicht sehr nützlich in Bezug auf die Eigenwerte, da diese in stark nichtlinearer Weise von der Matrix abhängen.

Dreiecksmatrix

Falls für $A \;=\; (\alpha_{ij}) \in \mathbb{R}^{n\times n}$

$$i > j \Rightarrow \alpha_{ij} = 0$$

gilt, so daß

$$A \;=\; \begin{pmatrix} \alpha_{11} & \cdots & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ & \cdots\cdots\cdots & & \\ 0 & \cdots & \cdots & \alpha_{nn} \end{pmatrix},$$

dann nennt man A eine **obere Dreiecksmatrix**.

Analog definiert man auch die **untere Dreiecksmatrix**, deren oberhalb der Hauptdiagonale stehenden Elemente Null sind.

Diagonale Matrizen

$A \in \mathbb{R}^{n\times n}$ heißt **diagonal**, wenn $i \neq j \Rightarrow \alpha_{i\,j} = 0$ gilt, also

$$A = \begin{pmatrix} \alpha_{1\,1} & 0 & \cdots & 0 \\ 0 & \alpha_{2\,2} & \cdots & 0 \\ & \ldots\ldots\ldots & & \\ 0 & 0 & \cdots & \alpha_{n\,n} \end{pmatrix}.$$

Man schreibt dann kurz $A = diag(\alpha_{i\,i})_{i=1\ldots n}$.

Insbesondere gilt

$$I = diag(1)_{i=1\ldots n}.$$

Summen und Produkte von diagonalen Matrizen sind wiederum diagonal:

$$\boxed{\begin{matrix} A = diag(\alpha_i)_{i=1\ldots n} \\ B = diag(\beta_i)_{i=1\ldots n} \end{matrix}} \quad \Longrightarrow \quad \boxed{\begin{matrix} A + B = diag(\alpha_i + \beta_i)_{i=1\ldots n} \\ A\,B = diag(\alpha_i\,\beta_i)_{i=1\ldots n} \end{matrix}}.$$

Orthogonale Matrizen

$A \in \mathbb{R}^{n\times n}$ heißt **orthogonal**, falls

$$A^T A = I = A\,A^T$$

wobei sich zeigen läßt, daß die zweite Identität aus der ersten folgt.
Bezeichnet man mit $\mathbf{a}_j = (\alpha_{i\,j})_{i=1\ldots n}$ den j-ten Spaltenvektor von A, so ist die Bedingung $A^T A = I$ äquivalent zu

$$\mathbf{a}_i \cdot \mathbf{a}_j = \begin{cases} 0 & \text{falls} \quad i \neq j \\ 1 & \text{falls} \quad i = j \end{cases}$$

Das heißt: Die Matrix A ist genau dann orthogonal, wenn ihre Spaltenvektoren eine orthonormale Basis von $\mathbb{R}^n$ bilden.

Da mit A auch A^T orthogonal ist, gilt dasselbe für die Zeilen von A, die ja die Spalten von A^T sind.

Produkt orthogonaler Matrizen

Für zwei orthogonale Matrizen A und B ist jeweils auch deren Produkt orthogonal, da

$$(AB)^T(AB) = (B^T A^T)(AB) = B^T(A^T A)B = B^T B = I.$$

Die Summe von orthogonalen Matrizen hat im allgemeinen nicht diese Eigenschaft. So ist zum Beispiel mit A auch $-A$ orthogonal, aber deren Summe, die Nullmatrix $A - A = 0$, sicherlich nicht.

Beispiel B.59 (Drehungen in der Ebene)

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \quad \Rightarrow \quad A^T = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

$$A^T A = \begin{pmatrix} \cos(\varphi)^2 + \sin(\varphi)^2 & \cos(\varphi)\sin(\varphi)\cdot(1-1) \\ \sin(\varphi)\cos(\varphi)\cdot(1-1) & \cos(\varphi)^2 + \sin(\varphi)^2 \end{pmatrix} = I$$

B - 7 Lösung linearer Gleichungssysteme

Lineare Systeme

Für eine lineare Abbildung

$$F : \mathcal{V} = Span\{\mathbf{v}_j\}_{j=1\dots n} \quad \to \quad \mathcal{W} = Span\{\mathbf{w}_i\}_{i=1\dots m}$$

und eine vorgegebene "Rechte Seite" $\mathbf{w} = \sum_{i=1}^m b_i \mathbf{w}_i$ mit $b_i \in \mathbb{R}$ findet man ein $\mathbf{v} = \sum_{j=1\dots n} x_j \mathbf{v}_j$ mit $F(\mathbf{v}) = \mathbf{w}$ durch Lösen des Gleichungssystems

$$\begin{array}{ccccccc}
\alpha_{1\,1}x_1 & + & \alpha_{1\,2}x_2 & + & \ldots + \alpha_{1\,j}x_j \ldots + \alpha_{1\,n}x_n & = & b_1 \\
\alpha_{2\,1}x_1 & + & \alpha_{2\,2}x_2 & + & \ldots + \alpha_{2\,j}x_j \ldots + \alpha_{2\,n}x_n & = & b_2 \\
 & & \multicolumn{3}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & & \\
\alpha_{i\,1}x_1 & + & \alpha_{i\,2}x_2 & + & \ldots + \alpha_{i\,j} \ldots + \alpha_{i\,n}x_n & = & b_i \\
 & & \multicolumn{3}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & & \\
\alpha_{m\,1}x_1 & + & \alpha_{m\,2}x_2 & + & \ldots + \alpha_{m\,j} \ldots + \alpha_{m\,n}x_n & = & b_m
\end{array}$$

Matrix–Vektor–Schreibweise

Äquivalenterweise ergibt sich in Matrix–Vektor–Schreibweise

$$A\mathbf{x} = \begin{pmatrix} \alpha_{1\,1} & \ldots & \alpha_{1\,j} & \ldots & \alpha_{1\,n} \\ \alpha_{2\,1} & \ldots & \alpha_{2\,j} & \ldots & \alpha_{2\,n} \\ & & \ldots\ldots\ldots\ldots & & \\ \alpha_{m\,1} & \ldots & \alpha_{m\,j} & \ldots & \alpha_{m\,n} \end{pmatrix} \mathbf{x} = \mathbf{b}$$

wobei $\mathbf{x} = (x_1, \ldots, x_j, \ldots, x_n)^T$ und $\mathbf{b} = (b_1, \ldots, b_i, \ldots, b_m)^T$ sind (unter Verletzung der Konvention, daß alle Skalare mit griechischen Buchstaben benannt sein sollten).

Man bezeichnet das lineare System von m Gleichungen in n Unbekannten als

unterbestimmt	wenn $m < n$
quadratisch	wenn $m = n$
überbestimmt	wenn $m > n$

Definition B.60 (Regularität)

Eine Abbildung $F : \mathbb{R}^n \to \mathbb{R}^n$ und entsprechende Matrizen A heißen regulär, falls

$$A\mathbf{x} = F(\mathbf{x}) = 0 \qquad \text{g.d.w.} \qquad \mathbf{x} = 0\,,$$

andernfalls heißen sie singulär.

Lemma B.61

Falls A regulär ist, dann hat $A\mathbf{x} = \mathbf{b}$ genau eine eindeutige Lösung für jedes $\mathbf{b}$.

Ein Kriterium, ob eine Matrix regulär oder singulär ist, liefert die im Abschnitt **B-9** eingeführte **Determinante** $\det(A)$.

Wünschenswerte Lösungsalgorithmen prüfen die Regularität und liefern entweder die eindeutige Lösung oder Singularitätsbeschreibungen.

Lösung Linearer Gleichungssysteme in Spezialfällen

Ist A eine Orthogonal-, Diagonal- oder Dreiecksmatrix (das sind diejenigen, deren Struktur sich auf das Produkt überträgt), so lassen sich die entsprechenden linearen Systeme $A\mathbf{x} = \mathbf{b}$ relativ leicht lösen.

Lemma B.62 (Lösung orthogonaler Systeme)

Falls A orthogonal ist, gilt:

$$A\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad A^T A\mathbf{x} = \mathbf{x} = A^T\mathbf{b}$$

In diesem Falle kann das Gleichungssystem also einfach durch die Multiplikation der rechten Seite $\mathbf{b}$ mit der Transponierten A^T gelöst werden.

Lemma B.63 (Lösung diagonaler Systeme)

Falls $A = \mathrm{diag}(\alpha_i)_{i=1\ldots n}$ eine Diagonalmatrix ist, so reduziert sich das lineare System auf die Gleichungen $\alpha_i x_i = b_i$. Diese werden für beliebige b_i durch $x_i = b_i/\alpha_i$ genau dann erfüllt, wenn keines der Diagonalelemente α_i gleich Null ist.

Falls diese Regularitätsbedingung verletzt ist, muß **b** *die Konsistenzbedingung*

$$\alpha_i = 0 \quad \Rightarrow \quad b_i = 0$$

erfüllen. Die entsprechenden Lösungskomponenten x_i sind dann beliebig, so daß das Gleichungssystem $A\mathbf{x} = \mathbf{b}$ mehrdeutig lösbar ist.

Lemma B.64 (Lösung von Dreieckssystemen)

Ist A eine untere Dreiecksmatrix, hat das entsprechende Gleichungssystem $A\mathbf{x} = \mathbf{b}$ die folgende "gestaffelte" Form:

$$\begin{array}{lcl}
\alpha_{1\,1}x_1 & = & b_1 \\
\alpha_{2\,1}x_1 + \alpha_{2\,2}x_2 & = & b_2 \\
\quad\vdots & \vdots & \\
\alpha_{i\,1}x_1 + \alpha_{i\,2}x_2 + \cdots + \alpha_{i\,i}x_i & = & b_i \\
\quad\vdots & \vdots & \\
\alpha_{n\,1}x_1 + \alpha_{n\,2}x_2 + \cdots + \alpha_{n,\,n-1}x_{n-1} + \alpha_{n\,n}x_n & = & b_n
\end{array}$$

Vorwärtssubstitution

Nun kann man zunächst aus der ersten Gleichung x_1 bestimmen, dann diesen Wert in die Zweite einsetzten, um x_2 zu erhalten, und so weiter. Unter der Regularitätsbedingung aus Lemma B.63, daß wiederum keines der diagonalen Elemente $\alpha_{i\,i}$ verschwindet, hat man also

$$\begin{array}{lcl}
x_1 & = & b_1/\alpha_{1\,1} \\
x_2 & = & (b_2 - \alpha_{2\,1}x_1)/\alpha_{2\,2} \\
x_3 & = & (b_3 - \alpha_{3\,1}x_1 - \alpha_{3\,2}x_2)/\alpha_{3\,3} \\
 & \vdots & \\
x_i & = & (b_i - \alpha_{i\,1}x_1 - \cdots - \alpha_{i\,i-1}x_{i-1})/\alpha_{i\,i} \\
 & \vdots & \\
x_n & = & (b_n - \alpha_{n\,1}x_1 - \cdots - \alpha_{n\,j}x_j - \cdots - \alpha_{n\,n-1}x_{n-1})/\alpha_{n\,n}
\end{array}$$

Man braucht $n(n-1)/2$ Multiplikationen und Additionen sowie n Divisionen.

Rückwärtssubstitution

Bei einer oberen Dreiecksmatrix A ergibt sich entsprechend das Verfahren der **Rückwärtssubstitution**, wobei jetzt die x_i für $i = n, n-1, \ldots, 1$ durch die Formel

$$x_i = \frac{1}{\alpha_{i\,i}}\left(b_i - \sum_{j=i+1}^{n} \alpha_{i\,j}x_j\right) \qquad\qquad i = n,\, n-1, \ldots, 1$$

bestimmt sind. Regularitätsbedingung ist wiederum, daß keines der Diagonalelemente verschwindet und der Rechenaufwand ist auch hier von der Ordnung $n^2/2$ arithmetische Operationen.

Zur Lösung allgemeiner linearer Systeme kann man die Matrix A so modifizieren, daß sie eine der oben genannten speziellen Formen annimmt oder das Produkt solcher spezieller Matrizen wird. Das klassische Verfahren für eine solche Transformation ist die **Elimination nach Carl Friedrich Gauß** (1777 – 1855).