Basistransformation

Definition: Wir bezeichnen mit M_n die Menge aller $n \times n-$Matrizen. Die Menge aller invertierbaren Matrizen (also aller $a \in M_n$, so daß es ein $A^{-1} \in M_n$ gibt, mit $A \cdot A^{-1} = A^{-1}A = E$) bezeichnen wir mit GL_n.

Bemerkung: Für eine $n \times n-$Matrix A gilt:

$$AE = (A\vec{e}_1, \ldots, A\vec{e}_n) = (\vec{a}_{\bullet 1}, \ldots, \vec{a}_{\bullet n})$$

wobei E die $n \times n$ Einheitsmatrix und die $\vec{e}_i$ die Standardbasisvektoren sind. Anders gesagt: Die Matrix ist durch ihr Verhalten auf der Standardbasis eindeutig bestimmt. Wir wollen nun untersuchen, was passiert, wenn wir von der Standardbasis auf eine andere Basis übergehen.

Erinnerung: i) Die Vektoren $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ bilden genau dann eine Basis, wenn die Matrix, deren Spalten die Vektoren $\vec{v}_i$ sind, invertierbar ist.

ii) Die Vektoren $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ heißen eine Basis von $\mathbb{R}$, wenn gilt:

$$(\vec{v}_1, \ldots, \vec{v}_n) \in GL_n.$$

Satz: Ist $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ und $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ eine Basis von $\mathbb{R}^n$, so ist

$$\vec{x} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$$

mit eindeutig bestimmten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Grund:

$$\vec{x} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \;\Leftrightarrow\; \vec{x} = (\vec{v}_1, \ldots, \vec{v}_n) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = (\vec{v}_1, \ldots, \vec{v}_n)^{-1} \cdot \vec{x}$$

□

Definition: Bezeichnen wir eine Basis mit $\mathcal{A}$ und die zugehörige Matrix, deren Spalten die Vektoren aus $\mathcal{A}$ sind mit $P_{\mathcal{A}}$, so ist

$$\vec{x}_{\mathcal{A}} := P_{\mathcal{A}}^{-1} \cdot \vec{x}$$

der Koeffizientenvektor von $\vec{x}$ bezüglich $\mathcal{A}$.

Motivation für die Definition: Da die Vektoren $\vec{a}_1, \ldots, \vec{a}_n$ eine Basis sind, gibt es für $\vec{x} \in \mathbb{R}^n$ eindeutig bestimmte $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, mit

$$\vec{x} = \lambda_1 \vec{a}_1 + \ldots + x_n \vec{\lambda}_n = (\ldots, \vec{a}_i, \ldots)\vec{\lambda}$$

also

$$\vec{\lambda} = (\ldots, \vec{a}_i, \ldots)^{-1}\vec{x}$$

.

Basiswechsel: Es ist für $A \in M_n$ und eine Basis $\mathcal{A}$ von $\mathbb{R}^n$:

$$\vec{y} = A\vec{x} \;\Leftrightarrow\; E\vec{y} = AE\vec{x} \;\Leftrightarrow\; P_{\mathcal{A}}P_{\mathcal{A}}^{-1}\vec{y} = AP_{\mathcal{A}}P_{\mathcal{A}}^{-1}\vec{x}$$

$$\Leftrightarrow\; P_{\mathcal{A}}\vec{y}_{\mathcal{A}} = AP_{\mathcal{A}}\vec{x}_{\mathcal{A}}$$

Die Matrix $P_{\mathcal{A}}^{-1}AP_{\mathcal{A}}$ transformiert also den Koeffizientenvektor $\vec{x}_{\mathcal{A}}$ in den Koeffizientenvektor $\vec{y}_{\mathcal{A}}$, wenn Die Matrix A den Vektor $\vec{x}$ in den Vektor $\vec{y}$ transformiert.

Definition:

$$M_{\mathcal{A}}(A) := P_{\mathcal{A}}^{-1}AP_{\mathcal{A}}$$

heißt Darstellungsmatrix von A bezüglich der Basis $\mathcal{A}$.

Es gilt:

$$M_{\mathcal{A}}(A) = \left(P_{\mathcal{A}}^{-1}\left(A\vec{v}_1\right), \ldots, P_{\mathcal{A}}^{-1}\left(A\vec{v}_n\right)\right)$$

wenn $P_{\mathcal{A}} = (\vec{v}_1, \ldots, \vec{v}_n)$ ist. In den Spalten von $M_{\mathcal{A}}(A)$ stehen also die Koeffizienten (bzgl. $\mathcal{A}$) der Bilder der Basisvektoren.

Merksatz: In den Spalten der Darstellungsmatrix stehen die Koeffizienten der Bilder der Basis vektoren.

Sei nun $\mathcal{B}$ eine weitere Basis von $\mathbb{R}^n$. Dann ist $M_{\mathcal{A}}(A) = P_{\mathcal{A}}^{-1}AP_{\mathcal{A}}$ und $M_{\mathcal{B}}(A) = P_{\mathcal{B}}^{-1}AP_{\mathcal{B}}$, also $A = P_{\mathcal{A}}M_{\mathcal{A}}(A)P_{\mathcal{A}}^{-1}$ und somit:

$$M_{\mathcal{B}}(A) = P_{\mathcal{B}}^{-1}P_{\mathcal{A}}M_{\mathcal{A}}P_{\mathcal{A}}^{-1}P_{\mathcal{B}} = (P_{\mathcal{A}}^{-1}P_{\mathcal{B}})^{-1}M_{\mathcal{A}}(P_{\mathcal{A}}^{-1}P_{\mathcal{B}})$$

Satz: Für zwei Matrizen $A, B \in M_n$ sind die folgenden beiden Eigenschaften äquivalent:

$$A \text{und } B \text{ sind ähnlich}$$

$$A \text{ und } B \text{ sind Darstellungsmatrizen derselben Matrix zu verschiedenen Basen}$$

Grund: Die Richtung $'' \Leftarrow ''$ haben wir oben gesehen. Um die Richtung $'' \Rightarrow ''$ einzusehen, wählen wir als Basis $\mathcal{C}$ die Spaltenvektoren von P. Dann gilt $M_{\mathcal{C}}(A) = B = M_{\mathcal{E}}(B)$, d.h. B ist die Darstellungsmatrix von A in der Basis $\mathcal{C}$ und von $\mathcal{B}$ zur Basis $\mathcal{E}$ (der Standardbasis).

Dies motiviert die folgende

Definition: Zwei Matrizen $A, B \in M_n$ heißen ähnlich, wenn es eine invertierbare Matrix P gibt, mit:

$$B = P^{-1}AP$$

Bemerkung: Diagonalmatrizen sind in mehrfacher Hinsicht besonders einfach.

Beispiele:

$$diag(\lambda_1, \dots, \lambda_n)^k = diag(\lambda_1^k, \dots, \lambda_n^k)$$

für alle $k \in \mathbb{Z}$.

Spezialfall:

$$diag(\lambda_1, \dots, \lambda_n)^{-1} = diag(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n})$$

falls alle $\lambda_i \neq 0$ sind.

Was ist die nächstbeste, naheliegende Eigenschaft? Ist $A \in M_n$ ähnlich zu einer Diagonalmatrix, so gilt:

$$A^k = P^{-1}DP \cdot P^{-1}DP \cdot \dots \cdot P^{-1}DP = P^{-1}D^kP$$

und

$$A^{-1} = P^{-1}D^{-1}P$$

Definition: Eine $n \times n-$Matrix heißt diagonalisierbar, wenn sie ähnlich zu einer Diagonalmatrix ist.

Bemerkung: Ist A diagonalisierbar, so gilt

$$AP^{-1} = P^{-1}D$$

Sind $\vec{p}_1, \dots, \vec{p}_n$ die Spaltenvektoren von P^{-1}, so heißt das:

$$(A\vec{p}_1, \dots, A\vec{p}_n) = (\lambda_1\vec{p}_1, \dots, \lambda_n\vec{p}_n)$$

für $D = diga(\lambda_1, \dots, \lambda_n)$ bzw.:

$$A\vec{p}_i = \lambda_i\vec{p}_i \text{ für alle} i = 1, \dots, n$$

Dies wird später die Begriffe Eigenvektoren und Eigenwerte begründen, also Vektoren $\neq \vec{0}$, die auf ein Vielfaches von sich selbst abgebildet werden.

Nicht alle Matrizen sind diagonalisierbar: Eine Drehung in der Ebene um den Winkel φ bildet Vektoren nicht auf Vielfaches von sich selbst ab (bis auf die Spezialfälle $\varphi = 0$ und $\varphi = \pi$).

Es stellt sich also die Frage, ob es eine hinreichend große Klasse von diagonalisierbaren Matrizen gibt. Das erfreuliche Ergebnis: Alle symmetrischen Matrizen sind diagonalisierbar.

Umgekehrt gilt:

Besitzt $\mathbb{R}^n$ eine Basis $\mathcal{A}$ bestehend aus Vektoren $\vec{v}_1, \ldots, \vec{v}_n$, so daß für die $n \times n-$ Matrix A gilt $A\vec{v}_i = \lambda_i \vec{v}_i$, so haben wir offensichtlich:

$$M_{\mathcal{A}}(A) = diag(\lambda_1, \ldots, \lambda_n)$$

Definition: Für eine $n \times n-$Matrix und einen Vektor $\vec{v} \neq \vec{0}$. Dann heißt $\vec{v}, A\vec{v}, \ldots, A^{k-1}\vec{v}$ eine Krylov-Sequenz von A und $\vec{v}$, wenn die Vektoren linear unabhängig sind und die Vektoren $\vec{v}, A\vec{v}, \ldots, A^{k-1}\vec{v}, A^k\vec{v}$ linear abhängig sind.

Wenn $k = n$ ist , bilden also $\vec{v}, \ldots, A^{n-1}\vec{v}$ eine Basis und dann ist

$$A\left(A^{n-1}\vec{v}\right) = \sum_{k=0}^{n-1} b_k A^k \vec{v}$$

Die Darstellungsmatrix von A bezüglich der obigen Basis ist also:

$$\begin{pmatrix} 0 & & & & b_0 \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & 0 & \\ & & & 1 & b_{n-1} \end{pmatrix}$$