Blockmatrizen

Ergänzung zum Kapitel 3 Vektoren und Matrizen

Dies ist ein Ergänzungsartikel zum Buch *Lineare Algebra: Eine anwendungsorientierte Einführung*, von Andreas Müller. Erschienen im Verlag Springer-Vieweg, ISBN 978-3-662-67865-7 und ISBN 978-3-662-67866-4 (eBook).

https://doi.org/10.1007/978-3-662-67866-4

Website zum Buch: https://linalg.ch

Springer-Link: https://link.springer.com/book/10.1007/978-3-662-67866-4

Blockmatrizen

Andreas Müller

Zusammenfassung

Die in Kapitel 3 eingeführte Matrizenalgebra funktioniert nicht nur, wenn die Matrixelemente Zahlen sind, sondern auch, wenn sie kleinere Matrizen sind. In diesem Artikel betrachten wir Matrizen als zusammengesetzt aus vielen kleineren Blöcken und drücken die wichtigsten Rechenoperationen für Matrizen durch Operationen mit den Blockmatrizen aus.

Es ist oft natürlich, eine Matrix in Blöcke zu zerlegen. Die beschleunigte Bewegung eines Massepunktes wird vollständig beschrieben durch einen Vektor, der die Ortskoordinaten, die Geschwindigkeitskoordinaten und die Beschleunigungen enthält. Aus diesen Größen zum Zeitpunkt t können die entsprechenden Größen zum Zeitpunkt $t+\Delta t$ bestimmt werden, indem man die Multiplikation

$$\left(\begin{array}{c} x(t+\Delta t)\\ y(t+\Delta t)\\ z(t+\Delta t)\\ \hline v_x(t+\Delta t)\\ v_y(t+\Delta t)\\ v_z(t+\Delta t)\\ \hline a_x(t+\Delta t)\\ a_y(t+\Delta t)\\ a_z(t+\Delta t)\end{array}\right) = \left(\begin{array}{ccc|ccc|ccc} 1&0&0&\Delta t&0&0&\frac12\Delta t^2&0&0\\ 0&1&0&0&\Delta t&0&0&\frac12\Delta t^2&0\\ 0&0&1&0&0&\Delta t&0&0&\frac12\Delta t^2\\ \hline 0&0&0&1&0&0&\Delta t&0&0\\ 0&0&0&0&1&0&0&\Delta t&0\\ 0&0&0&0&0&1&0&0&\Delta t\\ \hline 0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0&1\end{array}\right) \left(\begin{array}{c} x(t+\Delta t)\\ y(t+\Delta t)\\ z(t+\Delta t)\\ \hline v_x(t+\Delta t)\\ v_y(t+\Delta t)\\ v_z(t+\Delta t)\\ \hline a_x(t+\Delta t)\\ a_y(t+\Delta t)\\ a_z(t+\Delta t)\end{array}\right)$$

ausführt. Die Linien deuten bereits an, dass sich dies eleganter als eine Multiplikation einer Matrix von 3×-Blöcken mit Vektoren aus jeweils drei dreidimensionalen Vektoren schreiben lässt. Die kompaktere Form ist

$$\begin{pmatrix}\vec{s}(t+\Delta)\\ \vec{v}(t+\Delta)\\ \vec{a}(t+\Delta)\end{pmatrix} = \begin{pmatrix} I_3 & \Delta t\cdot I_3 & \frac12\Delta t^2\cdot I_3\\ 0 & I_3 & \Delta t\cdot I_3\\ 0&0&I_3\end{pmatrix}\begin{pmatrix} c\vec{s}(t)\\ \vec{v}(t)\\ \vec{a}(t)\end{pmatrix}.$$

Im Folgenden soll das Rechnen mit Blockmatrizen etwas vertieft werden. Unter anderem werden wir dabei eine Form des Gauß-Algorithmus für Blockmatrizen kennenlernen und daraus eine Formel für die Determinante ableiten sowie das Schur-Komplement definieren.

1 Blockmatrizen

Eine Matrix $A \in M_{n\times m}(\Bbbk)$ kann man sich auch als aufgeteilt in einzelne Blöcke vorstellen. Dazu braucht man Zahlen $n_1, \dots, n_k$ und $m_1, \dots, m_l$ mit

$$n_1 + \cdots + n_k = n \qquad \text{und} \qquad m_1 + \cdots + m_l = m.$$

Die Matrix zerfällt in $n_i \times m_j$ Blöcke A_{ij}:

$$A = \begin{array}{cc} & \begin{array}{cccc} m_1 & m_2 & & m_l \end{array} \\ \begin{array}{c} n_1 \\ n_2 \\ \\ n_k \end{array} & \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1l} \\ A_{21} & A_{22} & \cdots & A_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kl} \end{pmatrix} \end{array}$$

Wir nennen solche Matrix auch eine $(n_1, \dots, n_k) \times (m_1, \dots, m_l)$-Matrix.

Eine besonders regelmässige Form von Blockmatrizen entsteht beim Kronecker-Produkt von Abschnitt 15.3. Ist A eine $n \times m$-Matrix und B eine $p \times r$-Matrix, dann ist $A \otimes B$ eine Blockmatrix bestehend aus $m \times n$ Blöcken, die je die Abmessungen $p \times r$ haben. $A \otimes B$ ist also eine $(p, \dots, p) \times (r, \dots, r)$-Matrix, wobei im ersten Tupel n mal die Zahl p steht und im zweiten m mal die Zahl q.

2 Rechenoperationen

Die Rechenoperationen mit Matrizen kann man als Operationen mit den einzelnen Blöcken der Matrix schreiben, allerdings muss die Blockstruktur passen.

2.1 Addition

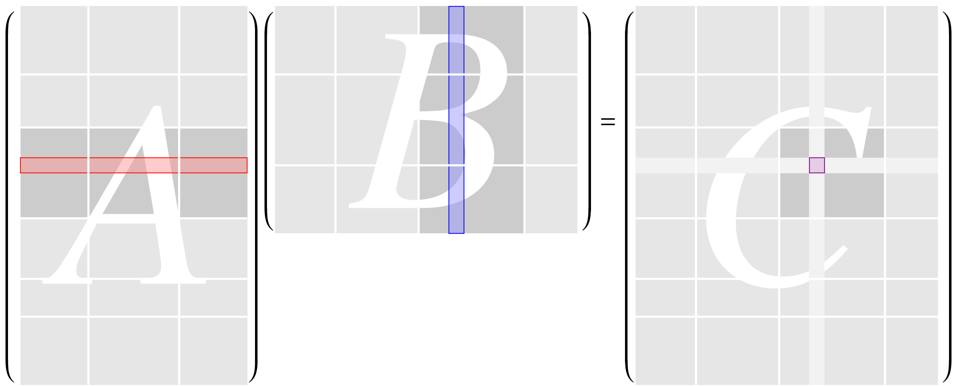
Um zwei Matrizen addieren zu können, müssen beide Matrizen die gleiche Blockstruktur haben. Wenn also A und B $(n_1, \dots, n_k) \times (m_1, \dots, m_l)$-Matrizen sind dann ist $C = A + B$ eine solche Matrix und die Blöcke von C sind

$$C_{ij} = A_{ij} + B_{ij}.$$

2.2 Multiplikation

Auch die Matrixmultipliation kann mit Blockmatrizen formuliert werden. Da das Matrizenprodukt als Produkt "Zeile×Spalte" ausgeführt wird, muss die Blockunterteilung der Zeilen des ersten Faktors zur Blockunterteilung der Spalten des zweiten Faktors passen. Sei also A eine $(n_1, \dots, n_k) \times (m_1, \dots, m_l)$-Matrix und B eine $(m_1, \dots, m_l) \times (r_1, \dots, r_s)$-Matrix, dann ist das Produkt $C = AB$ eine $(n_1, \dots, n_k) \times (r_1, \dots, r_s)$-Matrix.

Um einen Eintrag der Produktmatrix AB zu berechnen, sind alle Element der zugehörigen Zeile von A mit den Elementen der Spalte von B zu multiplizieren. Schematisch kann man diese Operation so darstellen:



Die Elemente der roten Zeile findet man also immer auf der gleichen Zeile in den dunkelgrauen Blöcken einer Zeile der Blockmatrix A, die blaue Spalte findet man immer in der gleichen Spalte in den dunkelgrauen Blöcken einer Spalte der Blockmatrix B. Das Produktelement steht im dunkelgrauen Block der Blockmatrix C auf der mit den Produkten der Zeilenblöcke von A und der Spalten blöcke von B nach

$$C_{ij} = \sum_{t=1}^{l} A_{it} B_{tj}$$

berechnet werden. Da A_{it} eine $n_i \times m_t$-Matrix ist und B_{tj} eine $m_t \times r_j$-Matrix, kann jedes einzelne Matrizenprodukt in der Summe ausgeführt werden.

2.3 Inverse Matrix

Ist A $(n_1, \dots, n_k) \times (n_1, \dots, n_k)$-Matrix, dann kann auch die inverse Matrix A^{-1} mit dieser Struktur interpretiert werden. Ist A zum Beispiel blockdiagonal mit Diagonalblöcken A_{ii}, dann ist auch A^{-1} blockdiagonal mit Diagonalblöcken A_{ii}^{-1}.

2.3.1 Dreiecksmatrix

Auch für eine obere Dreiecksmatrix kann man die Inverse direkt berechnen. Schreiben wir $B = A^{-1}$, dann muss

$$\begin{pmatrix} A_{11} & \dots & A_{1,k-2} & A_{1,k-1} & A_{1k} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & A_{k-2,k-2} & A_{k-2,k-1} & A_{k-2,k} \\ 0 & \dots & 0 & A_{k-1,k-1} & A_{k-1,k} \\ 0 & \dots & 0 & 0 & A_{kk} \end{pmatrix} \begin{pmatrix} B_{11} & \dots & B_{1,k-2} & B_{1,k-1} & B_{1k} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & B_{k-2,k-2} & B_{k-2,k-1} & B_{k-2,k} \\ 0 & \dots & 0 & B_{k-1,k-1} & B_{k-1,k} \\ 0 & \dots & 0 & 0 & B_{kk} \end{pmatrix} = \begin{pmatrix} I_{n_1} & \dots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & I_{n_{k-2}} & 0 & 0 \\ 0 & \dots & 0 & I_{n_{k-1}} & 0 \\ 0 & \dots & 0 & 0 & I_{n_k} \end{pmatrix}$$

Daraus kann man sofort ablesen, dass die Diagonalblöcke $B_{ii} = A_{ii}^{-1}$ sein müssen. Aus dem Produkt der zweitletzten Zeile von A mit der letzten Zeile von B kann man aber auch $B_{k-1,k}$ bestimmen:

$$A_{k-1,k-1}B_{k-1,k} + A_{k-1,k}B_{kk} = 0 \qquad \Rightarrow \qquad B_{k-1,k} = -A_{k-1,k-1}^{-1}A_{k-1,k}A_{kk}^{-1}$$

Entsprechend kann man mit nacheinander alle B_{ij} weiter oben und links bestimmen. Diese Vorgehensweise ist natürlich nichts anderes als das Rückwärtseinsetzen im Tableau

$$\left[\begin{array}{cccc|cccc} A_{11} & \dots & A_{1,k-1} & A_{1k} & I_{n_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & A_{k-1,k-1} & A_{k-1,k} & 0 & \dots & I_{n_{k-1}} & 0 \\ 0 & \dots & 0 & A_{kk} & 0 & \dots & 0 & I_{n_k} \end{array}\right]$$

Mit dieser Methode kann man zum Beispiel die Inversen der Dreiecksmatrizen

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -D^{-1}CA^{-1} & D^{-1} \end{pmatrix}$$

finden, die man auch durch direkte Rechnung verifizieren kann.

2.3.2 Der Fall einer $(p, q) \times (p, q)$-Matrix

Spezialfälle von $(p, q) \times (p, q)$-Matrizen haben wir schon in Abschnitt 2.3.1 kennengelernt, nämlich die Dreiecksblockmatrizen. Die Inverse einer beliebigen solchen Matrix aber ist uns noch nicht gelungen. Getreu dem Motto von Kapitel 12, dass die Faktorisierung einer Matrix ein Problem vereinfachen kann, versuchen wir die Inverse unter der Voraussetzung zu finden, dass wir eine Art LU-Zerlung von A in der Form

$$A = LU = \begin{pmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} I & U_{12} \\ 0 & I \end{pmatrix}$$

haben. Dann können wir die Inverse aus den Inversen der beiden Dreiecksteile finden, es ist

$$A^{-1} = U^{-1}L^{-1} = \begin{pmatrix} I & -U \\ 0 & I \end{pmatrix} \begin{pmatrix} L_{11}^{-1} & 0 \\ -L_{22}^{-1}L_{21}L_{11}^{-1} & L_{22}^{-1} \end{pmatrix} = \begin{pmatrix} L_{11}^{-1} + UL_{22}^{-1}L_{21}L_{11}^{-1} & -UL_{22}^{-1} \\ -L_{22}^{-1}L_{21}L_{11}^{-1} & L_{22}^{-1} \end{pmatrix}. \tag{1}$$

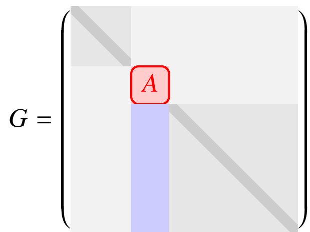
Sobald wir obige LU-Zerlegung finden können, können wir damit auch die Inverse der Blockmatrix berechnen.

3 Gauß-Algorithmus

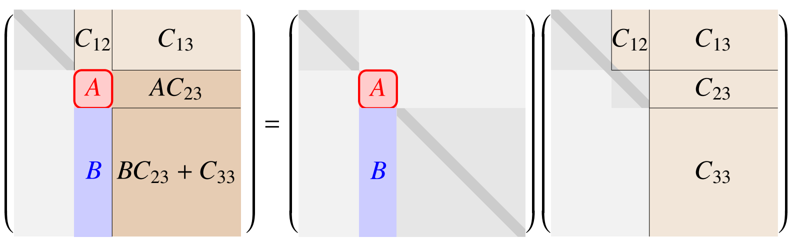
In der Formulierung als Matrixzerlegung faktorisiert der Gauß-Algorithmus eine Matrix A in eine linke untere Dreiecksmatrix L und eine unitale rechte obere Dreiecksmatrix U, also $A = LU$. In Abschnitt 12.2.1 wurden die Gauß-Matrizen eingeführt, mit denen sich die Schritt des Gauß-Algorithmus umkehren ließen. Dies ist die Form, in der wir den Gauß-Algorithmus auf Blockmatrizen verallgemeinern wollen.

3.1 Gauß-Operationen

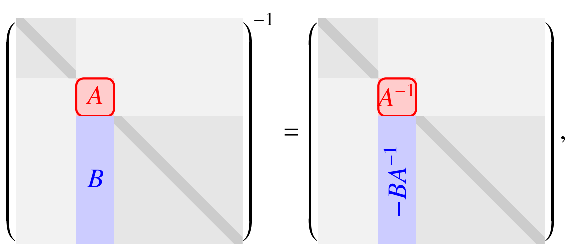
Im Gauß-Algorithmus wird durch Pivotelemente dividiert. Eine entsprechende Operation gibt es nur für quadratische Matrizen. Die Blockstrucktur der Matrix, mit der wir hier arbeiten wollen, muss daher einen quadrtischen Block als Pivot-Block enthalten. Einer Gauß-Matrix entsprechen daher Blockmatrizen der Form



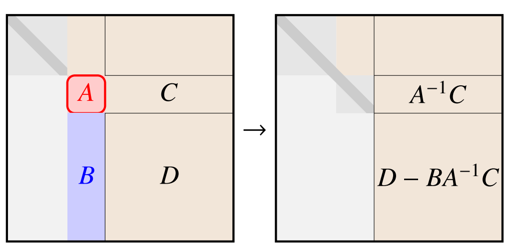
Damit können wir jetzt die Multiplikation mit der Gauß-Matrix rekonstruieren:



Die Inverse der Gauß-Matrix ist



wie man sich auch durch Nachrechnen überzeugen kann. Die Wirkung von G^{-1} bei der Multiplikation von links ist der Gauß-Schritt, denn wir für Tableaus wie folgt schreiben können:



Der Gauß-Algorithmus kann also ganz analog auf Blockmatrizen übertragen werden.

3.2 Der Fall einer $(p,q)\times(p,q)$-Matrix

Für eine $(p,q)\times(p,q)$-Matrix gibt es nur einen Gauß-Schritt zu tun und wir bekommen die LU-Zerlegung

$$\begin{aligned}\begin{pmatrix}A & B\\ C & D\end{pmatrix} &= \begin{pmatrix}A & 0\\ C & I\end{pmatrix}\begin{pmatrix}I & A^{-1}B\\ 0 & D-CA^{-1}B\end{pmatrix}\\ &= \begin{pmatrix}A & 0\\ C & I\end{pmatrix}\begin{pmatrix}I & 0\\ 0 & D-CA^{-1}B\end{pmatrix}\begin{pmatrix}I & A^{-1}B\\ 0 & I\end{pmatrix}\\ &= \underbrace{\begin{pmatrix}A & 0\\ C & D-CA^{-1}B\end{pmatrix}}_{=L}\underbrace{\begin{pmatrix}I & A^{-1}B\\ 0 & I\end{pmatrix}}_{=U}\end{aligned}$$

Aus der Formel (1) können wir jetzt die Inverse bestimmen:

$$\begin{pmatrix}A & B\\ C & D\end{pmatrix}^{-1} = \begin{pmatrix}A^{-1}+A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1}\\ -(D-CA^{-1}B)^{-1}CA^{-1} & (D-CA^{-1}B)^{-1}\end{pmatrix}$$

Für $p = q = 1$ sind die Einträge Zahlen und der gemeinsame Faktor $A^{-1}(D-CA^{-1}B)^{-1}$ kann ausgeklammert werden, was die Inverse

$$\begin{pmatrix}A & B\\ C & D\end{pmatrix}^{-1} = (AD-CB)^{-1}\begin{pmatrix}D & -B\\ -C & A\end{pmatrix},$$

ergibt, in Übereinstimmung mit der Formel (4.27).

4 Determinante

Für die Berechnung der Determinanten wurden im Kapitel 4 verschiedene Berechnungsmethoden entwickelt, die sich jetzt auch auf Blockmatrizen übertragen lassen.

4.1 Pivotproduktformel

Die Determinante von A konnte in Kapitel 4 mit dem Gauß-Algorithmus berechnet werden. Der Wert der Determinante war das Produkt der Pivot-Elemente. Der Gauß-Algorithus für Blockmatrizen zerlegt eine quadratische Matrix in ein Produkt von Gauß-Matrizen und unitale Blockmatrix. Die Determinaten von A kann mit der Produktformel für die Determinante berechnet werden. Seine G_i die Gauß-Matrizen und U die unitale Blockmatrix, dann wird

$$A = G_1\cdots G_k U.$$

Die Determinante ist dann

$$\det A = \det G_1 \cdot \ldots \cdot \det G_k \cdot\ \det U.$$

Als unitale Blockmatrix ist U auch eine unitale Matrix und hat daher die Determinanten $\det U = 1$. Ist A_i der Pivot-Block der Gauß-Matrix G_i, dann folgt

$$\det A = \det A_1 \cdot \ldots \cdot \det A_k.$$

Die Pivotproduktformel lässt sich also vollständig übertragen.

4.2 2×2-Blockmatrix

Für eine 2×2-Matrix ist die Determinante durch leicht zu merkende Formel

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{2}$$

gegeben. Diese Formel wurde mit Hilfe des Gauß-Algorithmus gefunden. Da wird den Gauß-Algorithmus für Blockmatrizen in Abschnitt 3 entwickelt haben, können wir jetzt eine entsprechende Formel für $(p,q)\times(p,q)$-Blockmatrizen finden. Aus der Faktorisierung

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ B & I \end{pmatrix} \begin{pmatrix} I & A^{-1}C \\ 0 & D - BA^{-1}C \end{pmatrix}$$

folgt für die Determinante

$$\det \begin{pmatrix} A & C \\ B & D \end{pmatrix} = \det A \cdot \det(D - BA^{-1}C).$$

Falls alle Blöcke quadratische Matrizen sind, falls also A und D die gleiche Dimension hat, lässt sich die Determinanten noch weiter zu

$$\det(AD - ABA^{-1}C) = \det(AD - (ABA^{-1})C)$$

vereinfachen. Die Formel sieht der Formel (2) sehr ähnlich, der einzige Unterschied ist, dass B durch ABA^{-1} ersetzt wird. Für den Fall, dass die Blöcke die Dimension 1×1 haben, ist $ABA^{-1} = B$ und es entsteht wieder die Formel (2).

5 Schur-Komplement

Mehrmals haben wir für bei der Bestimmung der LU-Zerlegung, der Inversen, oder bei der Berechnung der Determinanten einer $(p,q) \times (p,q)$-Matrix die Matrix $D - CA^{-B}$ angetroffen. Sie verdient daher einen eigenen Namen.

Definition 3.1 (Schur-Komplement)**.** *Das* Schur-Komplement *der Matrix A in der $(p,q) \times (p,q)$-Matrix*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

ist die Matrix $S = D - CA^{-1}B$.

Das Schur-Komplement ermöglicht jetzt, frühere Resultate etwas übersichtlicher zu schreiben. Zum Beispiel ist die Inverse

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix}$$

oder die Determinante

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \det A \cdot \det S.$$

Man kann das Schur-Komplement auch als die Reduktion der Grösse eines Gleichungssystems betrachten. Das Gleichungssystem in Blockform

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$$

ist ausgeschrieben

$$\begin{array}{ccccc} Ax & + & By & = & f \\ Cx & + & Dy & = & g \end{array}$$

Subtrahiert man das CA^{-1}-fache der ersten Zeile von der zweiten, erhält man

$$\begin{array}{cccc} Ax & + \quad By & = & f \\ & (D - CA^{-1}B)y & = & g - CA^{-1}f \end{array}$$

Die zweite Gleichung ist $Sy = g - CA^{-1}f$, woraus $y = S^{-1}(g - CA^{-1}f)$ folgt, die Inverse des Schur-Komplements erlaubt also y zu bestimmen. Damit ist y bekannt und x kann durch Auflösen der ersten Gleichung nach $x = A^{-1}(f - By)$ gefunden werden.