Calculus and linear algebra for biomedical engineering
Week 4: Inverse matrices and determinants

Hartmut Führ
fuehr@matha.rwth-aachen.de

Lehrstuhl A für Mathematik, RWTH Aachen

October 30, 2008

Overview

Motivation

Suppose that we have variables $x_1, \ldots, x_m, y_1, \ldots, y_n, z_1, \ldots, z_k$, which are related by the following linear equations:

$$\mathbf{x} = A\mathbf{y} \ , \ \mathbf{y} = B\mathbf{z} \ ,$$

with suitable matrices A, B.

Is there a way to directly express the dependence of $\mathbf{x}$ on $\mathbf{z}$ as $\mathbf{x} = C\mathbf{z}$, in terms of a matrix C?

The answer is yes, and the matrix C is the matrix product of A and B.

Matrix multiplication

Definition.

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times k}$. Denote the columns of B by $\mathbf{b}_1, \ldots, \mathbf{b}_k$. Then the matrix $AB \in \mathbb{R}^{m \times k}$ is defined as

$$AB = (A\mathbf{b}_1 | A\mathbf{b}_2 | \ldots | A\mathbf{b}_k) \ .$$

Hence the jth column of AB is obtained by multiplying A with the jth column of B.

Properties of matrix multiplication

Theorem. Let $A, A' \in \mathbb{R}^{m\times n}$ and $B, B' \in \mathbb{R}^{n\times k}$.

1. If $\mathbf{0}_{m,n}$ denotes the zero $m \times n$ matrix, then
$$\mathbf{0}_{k,m}A = \mathbf{0}_{k,n} \;,\; A\mathbf{0}_{n,k} = \mathbf{0}_{m,k} \;.$$
2. Distributive laws: $A(B + B') = AB + AB'$ and $(A + A')B = AB + A'B$.
3. Associativity: For all $C \in \mathbb{R}^{k\times \ell}$: $A(BC) = (AB)C$.
4. Note: Matrix multiplication is not commutative!

If we regard vectors $\mathbf{x} \in \mathbb{R}^n$ as $n \times 1$-matrices, matrix-by-vector multiplication is a special case of matrix-by-matrix multiplication. Recall motivation: If $\mathbf{x} = A\mathbf{y}$ and $\mathbf{y} = B\mathbf{z}$, how does $\mathbf{x}$ depend on $\mathbf{z}$? By associativity:

$$\mathbf{x} = A(B\mathbf{z}) = (AB)\mathbf{z}$$

Identity matrix

Definition. For $m \in \mathbb{N}$ let the ($m \times m$-) identity matrix I_m be defined by

$$I_m = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \\ & & \ddots & & & \\ & & & \ddots & & \\ 0 & \dots & \dots & 0 & 1 & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{m \times m} \; .$$

Then $I_m \mathbf{x} = \mathbf{x}$, for all $\mathbf{x} \in \mathbb{R}^m$.
As a consequence, for arbitrary $A \in \mathbb{R}^{n \times m}$:

$$A = I_n A = A I_m \; .$$

Inverse matrix

Definition. Let $A \in \mathbb{R}^{n \times n}$. Then $B \in \mathbb{R}^{n \times n}$ is called inverse of A, denoted $B = A^{-1}$, if

$$AB = BA = I_n \; .$$

If A has an inverse, A is called invertible or regular.

Theorem. Let $A, B \in \mathbb{R}^{n \times n}$.

(a) If $AB = I_n$ or $BA = I_n$, then $B = A^{-1}$.

(b) If A and B are invertible, then AB is invertible with $(AB)^{-1} = B^{-1}A^{-1}$.

Invertible matrices and linear systems of equations

Theorem. Let $A \in \mathbb{R}^{n\times n}$. Then the following are equivalent:

(a) A is invertible.

(b) For every vector $\mathbf{y} \in \mathbb{R}^n$, the system $A\mathbf{x} = \mathbf{y}$ has a unique solution.

(c) The linear system $A\mathbf{x} = \mathbf{0}$ has rank n.

Sketch of proof:

(a) $\Rightarrow$ (b): Define $\mathbf{x} = A^{-1}\mathbf{y}$. Then $A\mathbf{x} = A(A^{-1}\mathbf{y}) = I_n\mathbf{y} = \mathbf{y}$.

(b) $\Rightarrow$ (a): For $i = 1, \ldots, n$, find the solution $\mathbf{x}_i$ of

$$A\mathbf{x}_i = \mathbf{e}_i \ , \ \text{with } \mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$$

(entry 1 at ith position), and let $X = (\mathbf{x}_1|\ldots|\mathbf{x}_n)$. Then

$$AX = (A\mathbf{x}_1|\ldots|A\mathbf{x}_n) = (\mathbf{e}_1|\ldots|\mathbf{e}_n) = I_n \ .$$

Matrix inversion and systems of linear equations

Recall:

- Inverting a matrix $A \in \mathbb{R}^{n \times n}$ amounts to simultaneously solving n linear systems of equations $A\mathbf{x}_i = \mathbf{e}_i$.
  Here the coefficient matrix is always the same, only the right hand side changes.
- Linear systems are systematically solved by Gauss elimination. The steps in the Gauss elimination only depend on the coefficient matrix.

$\Rightarrow$ The Gauss algorithm should be useful for the computation of inverse matrices.

Matrix inversion via Gauss algorithm

Let $A \in \mathbb{R}^{n \times n}$ be given.

1. Introduce the extended matrix
$$\tilde{A} = (A|I_n) \in \mathbb{R}^{n \times 2n} \ .$$
2. Using Gauss elimination, transform the extended matrix to
$$\tilde{A}_1 = (B|C) \ ,$$
where $B \in \mathbb{R}^{n \times n}$ has upper triangular form. Make sure to apply all basic transformations to the right hand side also.
3. If B has rank $< n$, the matrix A is not invertible.
4. If B has rank n, apply basic transformations to $\tilde{A}_1$ to obtain
$$\tilde{A}_2 = (I_n|C') \ .$$
Then $C' = A^{-1}$.

Inverting a 4×4 matrix

We wish to invert A via the Gauss algorithm, where

$$A = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 3 & -2 & -3 & 4 \\ 4 & -1 & -3 & 10 \\ 2 & 0 & -5 & -4 \end{pmatrix}.$$

Following the usual scheme, we get

$$\begin{aligned} \tilde{A} &= \left(\begin{array}{cccc|cccc} 1 & -1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 3 & -2 & -3 & 4 & 0 & 1 & 0 & 0 \\ 4 & -1 & -3 & 10 & 0 & 0 & 1 & 0 \\ 2 & 0 & -5 & -4 & 0 & 0 & 0 & 1 \end{array}\right) \\ &\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & -1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 6 & -4 & 0 & 1 & 0 \\ 0 & 2 & -3 & -6 & -2 & 0 & 0 & 1 \end{array}\right) \end{aligned}$$

Inverting a 4×4 matrix

$$\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & -1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 5 & -3 & 1 & 0 \\ 0 & 0 & -3 & -8 & 4 & -2 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & -1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 5 & -3 & 1 & 0 \\ 0 & 0 & 0 & 1 & 19 & -11 & 3 & 1 \end{array}\right)$$

Now the entries above the diagonal are also removed using basic transformations

Inverting a 4 × 4 matrix

$$\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & -1 & -1 & 0 & -18 & 11 & -3 & -1 \\ 0 & 1 & 0 & 0 & -22 & 12 & -3 & -1 \\ 0 & 0 & 1 & 0 & -52 & 30 & -8 & -3 \\ 0 & 0 & 0 & 1 & 19 & -11 & 3 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -92 & 53 & -14 & -5 \\ 0 & 1 & 0 & 0 & -22 & 12 & -3 & -1 \\ 0 & 0 & 1 & 0 & -52 & 30 & -8 & -3 \\ 0 & 0 & 0 & 1 & 19 & -11 & 3 & 1 \end{array}\right)$$

Hence:

$$A^{-1} = \begin{pmatrix} -92 & 53 & -14 & -5 \\ -22 & 12 & -3 & -1 \\ -52 & 30 & -8 & -3 \\ 19 & -11 & 3 & 1 \end{pmatrix}$$

General definition of determinant

Definition. Let $A \in \mathbb{R}^{(n+1)\times(n+1)}$, and $1 \leq j \leq n+1$. Then the submatrix $A_j \in \mathbb{R}^{n\times n}$ is obtained by deleting the first column and jth row of A.

Definition. Let $A = (a_{i,j})_{i,j}$ be a matrix.

1. For $A = (a) \in \mathbb{R}^{1\times 1}$, we define $\det(A) = a$.
2. Let $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ be already defined. We then define $\det : \mathbb{R}^{(n+1)\times(n+1)} \to \mathbb{R}$ as follows:

$$\begin{aligned} \det(A) &= a_{1,1}\det(A_1) - a_{2,1}\det(A_2) \pm \ldots + (-1)^n a_{n+1,1}\det(A_{n+1}) \\ &= \sum_{j=1}^{n+1} (-1)^{j-1} a_{j,1} \det(A_j) \end{aligned}$$

Example: Computing a 2×2-determinant

Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

Then, in the notation of the definition, the submatrices A_1 and A_2 are given as

$$A_1 = (d) \in \mathbb{R}^{1\times 1}\ ,\ A_2 = (b) \in \mathbb{R}^{1\times 1}$$

and thus

$$\det(A) = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb\ .$$

Example: Computing a 3×3 determinant

Let

$$A = \begin{pmatrix} 1 & 7 & 4 \\ 3 & 3 & 2 \\ -1 & 6 & 5 \end{pmatrix} .$$

Using the definition, and the formula for 2×2 determinants,

$$\begin{aligned} \det(A) &= 1 \cdot \det \begin{pmatrix} 3 & 2 \\ 6 & 5 \end{pmatrix} - 3 \cdot \begin{pmatrix} 7 & 4 \\ 6 & 5 \end{pmatrix} \\ &+ (-1) \cdot \begin{pmatrix} 7 & 4 \\ 3 & 2 \end{pmatrix} \\ &= 1 \cdot (15 - 12) - 3(35 - 24) - 1(14 - 12) \\ &= -32 . \end{aligned}$$

Determinants and invertibility

Theorem. (Cramer's rule)
Let

$$A = (\mathbf{a}_1|\dots|\mathbf{a}_n) \in \mathbb{R}^{n\times n} ,$$

with $\det(A) \neq 0$. Let $\mathbf{y} \in \mathbb{R}^n$. For $i = 1, \dots, n$, define B_i by replacing the ith column of A by $\mathbf{y}$,

$$B_i = (\mathbf{a}_1|\dots|\mathbf{a}_{i-1}|\mathbf{y}|\mathbf{a}_{i+1}|\mathbf{a}_n) \in \mathbb{R}^{n\times n} .$$

Then the vector $\mathbf{x} = (x_1, \dots, x_n)^T$, with

$$x_i = \frac{\det(B_i)}{\det(A)} ,$$

is the unique solution of the linear equation $A\mathbf{x} = \mathbf{y}$.
Consequence. Let $A \in \mathbb{R}^{n\times n}$. Then A is invertible if and only if $\det(A) \neq 0$.

Example: Solving a 2×2 system with Cramer's rule

We want to solve the equation $\begin{pmatrix} 2 & 3 \\ 4 & 8 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 8 \\ -4 \end{pmatrix}$.

The coefficient matrix has determinant 4, thus there is a unique solution $\mathbf{x} = (x_1, x_2)^T$, with

$$x_1 = \frac{\det \begin{pmatrix} 8 & 3 \\ -4 & 8 \end{pmatrix}}{4} = \frac{76}{4} = 19$$

and

$$x_2 = \frac{\det \begin{pmatrix} 2 & 8 \\ 4 & -4 \end{pmatrix}}{4} = \frac{-40}{4} = -10$$

Note: For high dimensions, Cramer's rule is not an efficient way to solve linear systems!

Computational rules for determinants

Theorem. Let $A \in \mathbb{R}^{n\times n}$, $A = (a_{i,j})_{i,j=1\dots,n}$

(a) Let $A^T = (b_{i,j})_{i,j=1\dots n}$, where $b_{i,j} = a_{j,i}$. Then $\det(A^T) = \det(A)$.

(b) If $A = BC$ for suitable matrices $B, C \in \mathbb{R}^{n\times n}$, then $\det(A) = \det(B)\det(C)$.

(c) Let A be of the form

$$A = \begin{pmatrix} A_1 & B \\ \mathbf{0}_{n-k,k} & A_2 \end{pmatrix} ,$$

with $A_1 \in \mathbb{R}^{k\times k}$, $A_2 \in \mathbb{R}^{(n-k)\times(n-k)}$, and $C \in \mathbb{R}^{k\times(n-k)}$. Then

$$\det(A) = \det(A_1)\det(A_2) .$$

(d) If B is upper triangular with diagonal entries $d_1, \dots, d_n$, then $\det(B) = d_1 d_2 \dots d_n$.

Determinants via basic transformations

Theorem. Let $A, B \in \mathbb{R}^n$ be such that B is obtained from A by a basic transformation.

(a) If B is obtained by swapping two rows, then $\det(B) = -\det(A)$.

(b) If B is obtained by multiplication of a row with a constant c, then $\det(B) = c\det(A)$.

(c) If B is obtained by adding a multiple of a column to another column, then $\det(B) = \det(A)$.

The same rules apply if B is obtained by applying a basic transformation to columns of A.

Example: Using basic transformations

Recall the matrix A from above

$$A = \begin{pmatrix} 1 & 7 & 4 \\ 3 & 3 & 2 \\ -1 & 6 & 5 \end{pmatrix} .$$

Using basic transformations, we can compute $\det(A)$ as

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} 1 & 7 & 4 \\ 0 & -18 & -10 \\ 0 & 13 & 9 \end{pmatrix} \\ &= \det \begin{pmatrix} -18 & -10 \\ 13 & 9 \end{pmatrix} \\ &= -162 + 130 = -32 . \end{aligned}$$

Example: 4×4 determinant

Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 7 & 3 & 1 & 3 \\ 5 & 0 & 0 & 4 \\ 2 & 0 & 0 & 0 \end{pmatrix}$$

Then, after swapping rows

$$\det(A) = -\det \begin{pmatrix} 2 & 0 & 0 & 0 \\ 7 & 3 & 1 & 3 \\ 5 & 0 & 0 & 4 \\ 1 & 2 & 3 & 1 \end{pmatrix}$$

$$= -\det \begin{pmatrix} 2 & 7 & 5 & 1 \\ 0 & 3 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 3 & 4 & 1 \end{pmatrix} \quad (\text{ using } \det(B) = \det(B^T))$$

Example: 4×4 determinant

$$
\begin{aligned}
\ldots &= -2\det\begin{pmatrix} 3 & 0 & 2 \\ 1 & 0 & 3 \\ 3 & 4 & 1 \end{pmatrix} \quad \text{(definition of det)} \\
&= 2\det\begin{pmatrix} 0 & 3 & 2 \\ 0 & 1 & 3 \\ 4 & 3 & 1 \end{pmatrix} \quad \text{(swapping columns)} \\
&= 8\det\begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix} \quad \text{(definition of det)} \\
&= 8(9-2) = 56\ .
\end{aligned}
$$

Example: 4×4-determinant using block matrices

Again consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 7 & 3 & 1 & 3 \\ 5 & 0 & 0 & 4 \\ 2 & 0 & 0 & 0 \end{pmatrix}$$

Swapping columns transforms A into a block matrix:

$$\begin{aligned} \det(A) &= -\det \begin{pmatrix} 3 & 2 & 1 & 1 \\ 1 & 3 & 7 & 3 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 2 & 0 \end{pmatrix} \\ &= -\det \begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix} \det \begin{pmatrix} 5 & 4 \\ 2 & 0 \end{pmatrix} \\ &= -(9-2) \cdot (-8) = 56 \end{aligned}$$

Summary

- Matrix-by-matrix multiplication and its properties.
- Identity matrix
- Invertible matrices
- Computing inverse matrices by the Gauss algorithm
- Definition and basic properties of determinants
- Determinants and invertibility. Cramer's rule.
- Computing determinants: Gauss algorithm, block matrices, etc.