Topic(s): matrix inverses, multivariate Newton's method

Proposition 18.1.5. For any $n \times n$ matrix A, the following two conditions on A are equivalent:

- The linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Explicitly, for every output $\mathbf{b} \in \mathbb{R}^n$, there is a *unique* input $\mathbf{x} \in \mathbb{R}^n$ that solves the equation $A\mathbf{x} = \mathbf{b}$.
- There is an $n \times n$ matrix B for which $AB = I_n$ and $BA = I_n$. This means that $T_B : \mathbb{R}^n \to \mathbb{R}^n$ is inverse to T_A.

When these conditions hold, B is uniquely determined and is denoted A^{-1}.

We call any A satisfying these conditions **invertible**, and B is called the **inverse matrix** of A (A is the inverse matrix of B.)

Example 1. Verify that $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$ are matrix inverses of each other. Also, verify algebraically that $T_A \circ T_B$ and $T_B \circ T_A$ are the identity maps.

$AB = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$ and $BA = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$

Noting $T_A(x, y) = \begin{bmatrix} x+2y \\ 3x+5y \end{bmatrix}$ and $T_B(x,y) = \begin{bmatrix} -5x+2y \\ 3x-y \end{bmatrix}$, we compute

$$T_A \circ T_B(x, y) = T_A(-5x+2y, 3x-y) = \begin{bmatrix} (-5x+2y)+2(3x-y) \\ 3(-5x+2y)+5(3x-y) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

and

$$T_B \circ T_A(x, y) = T_B(x+2y, 3x+5y) = \begin{bmatrix} -5(x+2y)+2(3x+5y) \\ 3(x+2y)-(3x+5y) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Theorem 18.1.8. If A and B are $n \times n$ matrices that satisfy $AB = I_n$, then A is invertible and B is its matrix inverse.

If a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is such that $ad - bc \neq 0$, then its matrix inverse is given by

$$\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Example 2. Find the matrix inverse of $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Recall that this is the matrix associated with R_θ, the map that rotates counterclockwise by θ.

Since $ad - bc = (\cos\theta)(\cos\theta) - (-\sin\theta)(\sin\theta) = 1$,

$$(A_\theta)^{-1} = \frac{1}{1}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = A_{-\theta}.$$

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The **determinant** of A, denoted $\det(A)$, is $ad - bc$. Hence, the determinant of A being non-zero is equivalent to A being invertible. Although we will not deal with determinants of larger matrices in Math 51, it is true that non-zero determinant implies invertibility and vice-versa for $n \times n$ matrices.

Example 3. Find the inverse matrices of the following 2×2 matrices.

(a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

We get $\det(A) = 4 - 6 = -2$, and so, $A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$.

(b) $\begin{bmatrix} -6 & 2 \\ 3 & -1 \end{bmatrix}$ Since $\det(A) = 6 - 6 = 0$, A is not invertible.

$\begin{bmatrix} -6 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -6a+2c & -6b+2d \\ 3a-c & 3b-d \end{bmatrix}$ cannot be I_2. If $-6b+2d=0$, then $3b-d=0$.

Example 4. Solve the following system of linear equations

$$\begin{aligned} x + 2y &= 7 \\ 3x + 4y &= 19 \end{aligned}$$

The system can be written as $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 19 \end{bmatrix}$. Multiplying both sides on the left by the inverse found above, we get

$$\begin{bmatrix} x \\ y \end{bmatrix} = I_2\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 7 \\ 19 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

Hence, $x = 5$ and $y = 1$.

Example 5. Solve the following system of linear equations

$$\begin{aligned} 2x + y - z &= 4 \\ x - 2y + z &= 3 \\ x - y - z &= 1 \end{aligned}$$

by using the fact that the inverse of $\begin{bmatrix} 2 & 1 & -1 \\ 1 & -2 & 1 \\ 1 & -1 & -1 \end{bmatrix}$ is $\frac{1}{7}\begin{bmatrix} 3 & 2 & -1 \\ 2 & -1 & -3 \\ 1 & 3 & -5 \end{bmatrix}$.

Writing the system as $\begin{bmatrix} 2 & 1 & -1 \\ 1 & -2 & 1 \\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 1 \end{bmatrix}$ and multiplying both sides on the left side by the inverse of the coefficient matrix gives us

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 3 & 2 & -1 \\ 2 & -1 & -3 \\ 1 & 3 & -5 \end{bmatrix}\begin{bmatrix} 2 & 1 & -1 \\ 1 & -2 & 1 \\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 3 & 2 & -1 \\ 2 & -1 & -3 \\ 1 & 3 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 1 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 17 \\ 2 \\ 8 \end{bmatrix}.$$

Thus, $x = \frac{17}{7}$, $y = \frac{2}{7}$, $z = \frac{8}{7}$.

Example 6. Consider the systems of linear equations

$$\begin{aligned} x + 3y - z &= 4 \\ 2x + y + 3z &= 3 \\ -x + y - 3z &= 1 \end{aligned}$$

and

$$\begin{aligned} x + 3y - z &= 5 \\ 2x + y + 3z &= 5 \\ -x + y - 3z &= -1 \end{aligned}$$

It turns out that $\begin{bmatrix} 1 & 3 & -1 \\ 2 & 1 & 3 \\ -1 & 1 & -3 \end{bmatrix}$ is not invertible.

Since both systems have the same coefficient matrix, we can apply the same operations to eliminate x:

① + ③	$4y - 4z = 5$	$4y - 4z = 4$
② + 2③	$3y - 3z = 5$	$3y - 3z = 3$

The two equations from the purple system are inconsistent, i.e. there are no solutions. The two equations from the green system are the same, i.e. there are infinitely many solutions. More specifically, if we set $z = t$, we get $y = t + 1$ and $x = -2t + 2$. Hence,

$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ is a solution for all $t \in \mathbb{R}$.

Theorem 18.3.3. Let A be an $n \times n$ matrix and consider the system of n equations in n variables $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is a given n-vector.

- If A is invertible, then this system has a *unique* solution, namely $\mathbf{x} = A^{-1}\mathbf{b}$. In particular, $A\mathbf{x} = \mathbf{0}$ has $\mathbf{x} = \mathbf{0}$ as its *only* solution when A is invertible.
- If A is *not* invertible, then $A\mathbf{x} = \mathbf{b}$ either has no solution or has infinitely many solutions. In particular, $A\mathbf{x} = \mathbf{0}$ has (infinitely many) non-zero solutions.

Thus, A is invertible precisely when $A\mathbf{x} = \mathbf{0}$ has $\mathbf{x} = \mathbf{0}$ as its *only* solution, and A is not invertible precisely when $A\mathbf{x} = \mathbf{0}$ has a non-zero solution.

Example 7. Recall, from Lecture 14, Example 4, that $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ stretches horizontally by a factor of 2 and $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ rotates 90 degrees clockwise.

So, we saw that $AB = \begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix}$ first rotates 90 degrees clockwise then stretches horizontally by a factor of 2.

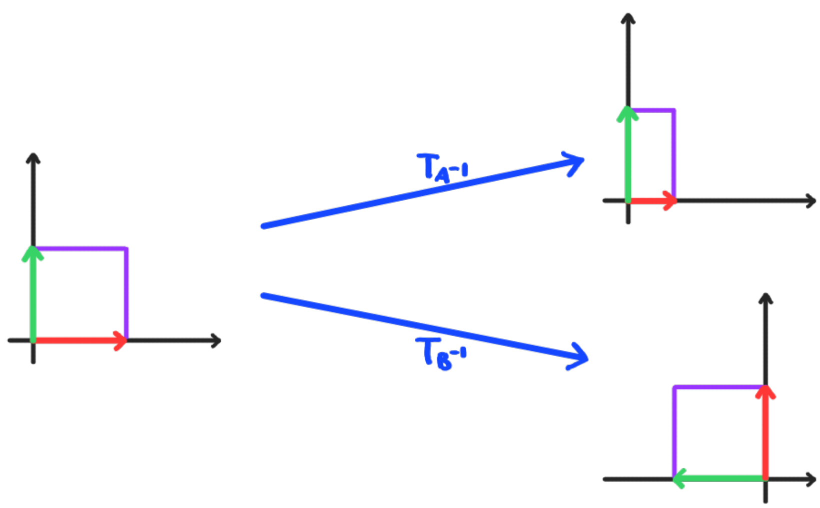
(a) Compute A^{-1} and B^{-1} algebraically.

We compute

$$A^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B^{-1} = \frac{1}{1}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

(b) Compute A^{-1} and B^{-1} geometrically.

A^{-1} compresses horizontally by 2 and B^{-1} rotates 90° counterclockwise.



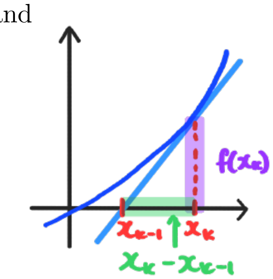
So, $A^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$ and $B^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

(c) Compute $(AB)^{-1}$. What is it in terms of A^{-1} and B^{-1}?

We see that $(AB)^{-1} = \frac{1}{2}\begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} = B^{-1}A^{-1}$.

It is always true that $(AB)^{-1} = B^{-1}A^{-1}$.

Newton's method for approximating zeros of non-linear functions. Let $\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^n$ be a non-linear function and let $\mathbf{a}$ be an initial guess for a solution to $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ with $D\mathbf{f}(\mathbf{a})$ invertible. Then, if we have chosen $\mathbf{a}$ reasonably, the sequence of vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots \in \mathbb{R}^n$ defined by $\mathbf{a}_1 = \mathbf{a}$ and



$$\vec{f}(\vec{a}_k) = D\vec{f}(\vec{a}_k)(\vec{a}_k - \vec{a}_{k-1}) \longleftrightarrow \mathbf{a}_{k+1} = \mathbf{a}_k - (D\mathbf{f}(\mathbf{a}_k))^{-1}\mathbf{f}(\mathbf{a}_k)$$

makes sense and converges rapidly to a solution of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

$f(x_k) = f'(x_k)(x_k - x_{k-1})$

Example 8. Define $\mathbf{f}: \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\mathbf{f}(x, y) = (x^2 + 2y - 2, x^3 - 2xy + 1).$$

This vanishes at approximately $(-1.19149, 0.290178)$.

(a) Compute $D\mathbf{f}(x, y)$ symbolically and use this to give a symbolic expression in terms of matrices and vectors for the output of one step of Newton's method.

We see that $D\vec{f}(x,y) = \begin{bmatrix} 2x & 2 \\ 3x^2 - 2y & -2x \end{bmatrix}$, and so, one step of Newton's method gives

$$\begin{bmatrix} x \\ y \end{bmatrix} - \left(\begin{bmatrix} 2x & 2 \\ 3x^2 - 2y & -2x \end{bmatrix}\right)^{-1} \begin{bmatrix} x^2 + 2y - 2 \\ x^3 - 2xy + 1 \end{bmatrix}.$$

(b) Compute the output of one step of Newton's method for $\mathbf{f}$ on the input $(2, 0)$.

Since $D\vec{f}(2,0) = \begin{bmatrix} 4 & 2 \\ 12 & -4 \end{bmatrix}$, we get $(D\vec{f}(2,0))^{-1} = \frac{1}{-40}\begin{bmatrix} -4 & -2 \\ -12 & 4 \end{bmatrix}$ and

$$\begin{bmatrix} 2 \\ 0 \end{bmatrix} - \frac{1}{-40}\begin{bmatrix} -4 & -2 \\ -12 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 9 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \frac{1}{40}\begin{bmatrix} -26 \\ 12 \end{bmatrix} = \begin{bmatrix} \frac{27}{20} \\ \frac{3}{10} \end{bmatrix}.$$

(c) Compute the output of one step of Newton's method for $\mathbf{f}$ on the input $(2, 2)$.

Since $D\vec{f}(2,2) = \begin{bmatrix} 4 & 2 \\ 8 & -4 \end{bmatrix}$, we get $(D\vec{f}(2,2))^{-1} = \frac{1}{-32}\begin{bmatrix} -4 & -2 \\ -8 & 4 \end{bmatrix}$ and

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} - \frac{1}{-32}\begin{bmatrix} -4 & -2 \\ -8 & 4 \end{bmatrix}\begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \frac{1}{32}\begin{bmatrix} -26 \\ -44 \end{bmatrix} = \begin{bmatrix} \frac{19}{16} \\ \frac{5}{8} \end{bmatrix}.$$

(d) What happens if you try to run a step of Newton's method for $\mathbf{f}$ on the input $(2, 10)$?

We cannot apply Newton's method since $D\vec{f}(2,10) = \begin{bmatrix} 4 & 2 \\ -8 & -4 \end{bmatrix}$ is not invertible.