Inverses of Square Matrices

A. Havens

Department of Mathematics
University of Massachusetts, Amherst

February 23-26, 2018

Outline

Inverses for Products: Division

- For nonzero real (or rational, or complex) numbers, there is a notion of multiplicative inverse, related to the multiplicative identity.
- Given a number $a \neq 0$, the *multiplicative inverse of a* is the number denoted a^{-1} with the property that

$$aa^{-1} = a^{-1}a = 1\,.$$

- 1 serves as the identity for multiplication, since $1 \cdot a = a = a \cdot 1$ for any a (either rational, real, or complex).
- For nonzero rational numbers p/q, the inverse is just the *reciprocal* q/p. For complex numbers $a + bi$, the inverse is given by the expression

$$(a+bi)^{-1} = \frac{a - bi}{a^2 + b^2}\,.$$

Inverses for Functions - Left and Right

In analogy with the number systems discussed above, we can define inverses for functions. However, there are some subtleties that arise when the domain and codomain are different sets.

Definition

A function $f : X \to Y$ is said to be *left invertible* if there exists a function $g : f(Y) \to X$ such that $g\big(f(x)\big) = x$ for every $x \in X$, i.e., $g \circ f = \mathrm{Id}_X$ (the identity function on X).

A function $f : X \to Y$ is said to be *right invertible* if there exists a function $h : Y \to X$ such that $f\big(h(y)\big) = y$, i.e., $f \circ h = \mathrm{Id}_Y$ (the identity function on Y).

Conditions for Left and Right Invertibility

- Note that if $f(x_1) = f(x_2) = y$ for $x_1 \neq x_2$, then there can be no function g such that $g \circ f = \mathrm{Id}_X$, as one cannot determine which value, x_1 or x_2, to assign to this y.
- Thus, left invertibility requires *injectivity*, i.e. f must be one-to-one. Conversely, an injective function always has a left-inverse. You should convince yourself that the left inverse function g must be surjective.
- Right invertibility requires *surjectivity* (f must be onto), and surjective functions always have right inverses. You should convince yourself that the right inverse function h must be injective.

Invertible $\iff$ One-To-One and Onto

- To have both a left and right inverse, a function must be both injective and surjective. Such functions are called bijective.
- Bijective functions always have both left and right inverses, and are thus said to be *invertible*.
- A function which fails to be either injective or surjective will fail to have either a left or right inverse, respectively.
- It is a good exercise in understanding functions and the definitions of injectivity and surjectivity to convince yourself of the above facts, by completing the **Challenge Problem**: prove that both left and right inverses of a function $f : X \to Y$ exist if and only if f is both one-to-one and onto.

Invertible $\iff$ One-To-One and Onto

In the special case of a transformation $f : X \to X$ which is both one-to-one and onto, (i.e. a bijective transformation of X), we can define a *total inverse*.

Definition

The *total inverse* of a bijective function $f : X \to X$ is a function denoted f^{-1} such that

$$f \circ f^{-1} = \mathrm{Id}_X = f^{-1} \circ f\,.$$

Example

For any of the number systems $\Bbbk = \mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, and fixed nonzero $\alpha \in \Bbbk$, the multiplication map $\mu_\alpha : \Bbbk \to \Bbbk$ given by $\mu_\alpha(x) = \alpha x$ is bijective, and its total inverse is $\mu_\alpha^{-1} = \mu_{\alpha^{-1}} = \mu_{1/\alpha}$.

Invertible Square Matrices

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ may be represented by a matrix $\mathrm{A} \in \mathbb{R}^{n\times n}$.

If the transformation T is invertible, then there exists a map $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ such that $T\big(T^{-1}(\mathbf{x})\big) = \mathbf{x} = T^{-1}\big(T(\mathbf{x})\big)$ for any $\mathbf{x} \in \mathbb{R}^n$. This map also has a representation as a matrix-vector product, say $T(\mathbf{x}) = \mathrm{C}\mathbf{x}$ for a matrix $\mathrm{C} \in \mathbb{R}^{n\times n}$.

Definition

A square matrix $\mathrm{A} \in \mathbb{R}^{n\times n}$ is said to be *invertible* if there exists a matrix $\mathrm{C} \in \mathbb{R}^{n\times n}$ such that

$$\mathrm{CA} = \mathrm{I}_n = \mathrm{AC}\,.$$

If no inverse exists, A is said to be *singular*. A square matrix which is *non-singular* is invertible.

Uniqueness

If A is invertible, then the inverse matrix C is unique.

Indeed, suppose that C and B are both inverses, so

$$\mathrm{BA} = \mathrm{I}_n = \mathrm{AB} \quad \text{and} \quad \mathrm{CA} = \mathrm{I}_n = \mathrm{AC}\,.$$

Then

$$\mathrm{B} = \mathrm{BI}_n = \mathrm{B}(\mathrm{AC}) = (\mathrm{BA})\mathrm{C} = \mathrm{I}_n\mathrm{C} = \mathrm{C}\,.$$

Thus, if A is invertible we will denote the unique inverse by $\mathrm{C} =: \mathrm{A}^{-1}$.

Properties

Theorem

Let $\mathrm{A}, \mathrm{B} \in \mathbb{R}^{n\times n}$ *be non-singular matrices, and* A^{-1}, B^{-1} *their respective inverses. Then the following properties hold:*

- *The maps* $\mathbf{x} \mapsto \mathrm{A}\mathbf{x}$ *and* $\mathbf{x} \mapsto \mathrm{A}^{-1}\mathbf{x}$ *are one-to-one and onto, whence* $\mathrm{RREF}(\mathrm{A}) = \mathrm{I}_n = \mathrm{RREF}(\mathrm{A}^{-1})$
- *The inverse matrix* A^{-1} *is also non-singular, and the inverse of the inverse is the original matrix:* $(\mathrm{A}^{-1})^{-1} = \mathrm{A}$.
- *The inverse of a product is the product of inverses, in opposite order:* $(\mathrm{AB})^{-1} = \mathrm{B}^{-1}\mathrm{A}^{-1}$
- *The transpose of the inverse is the inverse of the transpose:* $(\mathrm{A}^{-1})^{\mathrm{t}} = (\mathrm{A}^{\mathrm{t}})^{-1}$

The 2×2 Inverse Matrix Formula

Theorem

Let $\mathrm{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ for real numbers a, b, c and d. Then A is non-singular if and only if $ad - bc \neq 0$. In the case that A is non-singular,

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

There are a number of ways to arrive at this result, but the core details were proven in our previous discussion of the complete solution of 2×2 systems.

An Example

Example

The inverse of the matrix $\begin{bmatrix} 4 & -6 \\ -1 & 2 \end{bmatrix}$ is

$$\begin{bmatrix} 4 & -6 \\ -1 & 2 \end{bmatrix}^{-1} = \frac{1}{4(2)-(-6)(-1)}\begin{bmatrix} 2 & 6 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1/2 & 2 \end{bmatrix}.$$

Next, we'll use 2×2 matrices to represent and describe complex arithmetic, giving a geometric interpretation to complex multiplication.

Representing Complex Arithmetic with Matrices

Observe that the right-angle rotation matrix $\mathrm{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ squares to the negative of the 2×2 identity matrix.

This is analogous to the imaginary unit i, whose square is -1.

One can represent a complex number by matrices:

$$a + bi \sim a\mathrm{I}_2 + b\mathrm{J} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

You should check that the matrix product $(a\mathrm{I}_2 + b\mathrm{J})(c\mathrm{I}_2 + d\mathrm{J})$ gives a matrix representing the complex product $(a + bi)(c + di)$. You should note that these matrices commute in the product!

Representing Complex Arithmetic with Matrices

These matrix representatives of complex numbers add, scale and multiply just as complex numbers do. But what about inverses? Applying the formula for the 2×2 inverse, we get

$$\begin{aligned}(a\mathrm{I}_2 + b\mathrm{J})^{-1} &= \frac{1}{a(a) - (-b)(b)} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \\ &= \frac{1}{a^2 + b^2} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \sim \frac{a - bi}{a^2 + b^2} = \frac{1}{a + bi} \,.\end{aligned}$$

Thus, the arithmetic of $\mathbb{C}$ is identical to that of matrices of the form $a\mathrm{I}_2 + b\mathrm{J}$ for real scalars a and b, with matrix addition, products, and real scaling taking the roles of complex number addition, multiplication, and scaling by a real number.

Representing Complex Arithmetic with Matrices

Let $a + bi$ be a unit complex number. Then $a^2 + b^2 = 1$, so $a + bi$ represents a point on the unit circle, whence there is some angle θ such that $a = \cos\theta$ and $b = \sin\theta$.

Then the matrix representing $a + bi$ is precisely a rotation matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}!$$

Thus, multiplication by a unit complex number corresponds geometrically to a rotation!

In this way, the circle in the complex plane parametrizes the rotations of the plane.

Representing Complex Arithmetic with Matrices

For a fixed $\alpha = a + bi \in \mathbb{C}$, the map $\mu_\alpha : \mathbb{C} \to \mathbb{C}$ taking a generic complex number $z = x + yi$ to $\mu_\alpha(z) = \alpha z$ may be represented by the linear map of $\mathbb{R}^2$:

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \mapsto r \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

where $r = |\alpha| = \sqrt{a^2 + b^2}$ is the *modulus*, and θ is the *argument*, i.e., the angle between α and 1 in $\mathbb{C}$. This corresponds to the *polar* representation of a complex number: $\alpha = re^{i\theta} = r(\cos\theta + i\sin\theta)$. Thus general complex multiplication acts geometrically by an orientable similarity transformation, namely, as a rotation plus a dilation/contraction, dependent on the argument θ and the modulus $\sqrt{a^2 + b^2}$.

Conditions for Existence

Consider a linear transformation $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ of $\mathbb{R}^n$. What must hold about A for the inverse to exist?

Since the transformation must be bijective to be invertible, we know that the columns of A must be linearly independent and span $\mathbb{R}^n$. Therefore it must have n pivots, i.e., $\mathrm{RREF}(\mathrm{A}) = \mathrm{I}_n$.

A System approach

Let $\mathrm{A} \in \mathbb{R}^{n\times n}$ be a non-singular matrix with inverse A^{-1}. Write $\mathrm{A} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ and $\mathrm{A}^{-1} = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$.

Since $\mathrm{A}\mathrm{A}^{-1} = \mathrm{I}_n$, by the definition of matrix products

$$\begin{bmatrix} \mathrm{A}\mathbf{v}_1 & \dots & \mathrm{A}\mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix}.$$

Treating the components of $\mathbf{v}_i$ as variables, we get n systems of n equations in n unknowns, each of the form $\mathrm{A}\mathbf{v}_i = \mathbf{e}_i$. Since $\mathrm{RREF}(\mathrm{A}) = \mathrm{I}_n$, we can solve these by reduction.

Computing the Inverse of a Square Matrix

The row operations depend only upon A, so we may perform the reduction for all $\mathbf{v}_i$ simultaneously by forming the augmented matrix $\left[\mathrm{A} \mid \mathrm{I}_n\right]$, and performing Gauss-Jordan.

This leads to the following algorithm to determine if a given square matrix $\mathrm{A} \in \mathbb{R}^{n\times n}$ is non-singular and to find its inverse when it is:

1. Form the augmented matrix $\mathrm{M} = \left[\mathrm{A} \mid \mathrm{I}_n\right]$
2. Row-reduce by Gauss-Jordan.
3. If M reduces to the form $\left[\mathrm{I}_n \mid \mathrm{C}\right]$ for some matrix C, then $\mathrm{A}^{-1} = \mathrm{C}$. Otherwise, A is singular.

Computing the Inverse of a Square Matrix: An Example

Example

Find

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & -2 & 1 \end{bmatrix}^{-1}.$$

Form the augmented matrix

$$\left[\begin{array}{ccc|ccc} 0 & 1 & -1 & 1 & 0 & 0 \\ -1 & 0 & 2 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right].$$

What row operations are needed to reduce this?

Computing the Inverse of a Square Matrix: An Example

Example

Performing the row operations $R_2 \leftrightarrow R_1$, $-R_1 \mapsto R_1$, $R_3 - R_1 + 2R_2$, $R_2 + R_3 \mapsto R_2$, and $R_1 + 2R_3 \mapsto R_1$ yields the row equivalence

$$\left[\begin{array}{ccc|ccc} 0 & 1 & -1 & 1 & 0 & 0 \\ -1 & 0 & 2 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & 1 & 2 \\ 0 & 1 & 0 & 3 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 1 \end{array}\right].$$

Thus

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & -2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 4 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

One can check by multiplication that this is indeed the inverse.

Solving a System with an Inverse

If a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is given by $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ for a non-singular matrix A, then for any $\mathbf{b} \in \mathbb{R}^n$ the linear system $\mathrm{A}\mathbf{x} = \mathbf{b}$ has a unique solution.

Indeed, observe that

$$\mathrm{A}^{-1}(\mathrm{A}\mathbf{x}) = \mathrm{A}^{-1}\mathbf{b} \implies \mathbf{x} = \mathrm{A}^{-1}\mathbf{b}\,,$$

and $\mathbf{x} = \mathrm{A}^{-1}\mathbf{b}$ is indeed a solution since

$$\mathrm{A}(\mathrm{A}^{-1}\mathbf{b}) = \mathrm{I}_n\mathbf{b} = \mathbf{b}\,.$$

In practice, it is more efficient to solve by row reduction directly than to compute A^{-1}. However, there are some applications where the inverse is meaningful and useful to know, especially if one is interested in the map T^{-1} for geometric reasons.

An Example

Example

Use the inverse matrix to solve the system

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}.$$

Since we already computed the inverse, we can simply multiply to solve:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 2 \end{bmatrix}.$$

Another Way to Reflect About a Line in $\mathbb{R}^2$

Last time, we saw that if a line $\ell = \operatorname{Span}\mathbf{v}$, then the reflection through this line was the map

$$\mathbf{x} \mapsto 2\frac{\mathbf{v}\cdot\mathbf{x}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} - \mathbf{x} = (2\mathrm{proj}_{\mathbf{v}} - \mathrm{I}_2)\mathbf{x}\,.$$

This allows one to describe a matrix for reflection through ℓ. Here's another way to arrive at a formula for such a matrix:

- Rotate the line ℓ to coincide with the x_1 axis (so $\mathbf{v}$ is rotated to be in the same direction as $\mathbf{e}_1$.)
- Then reflect through the x_1 axis.
- Then rotate back so that $\mathbf{e}_1$ rotates to be in the same direction as $\mathbf{v}$.

Another Way to Reflect About a Line in $\mathbb{R}^2$

Suppose **v** makes an angle of θ with the x_1-axis.
Let R_θ represent rotation by the angle θ, and let

$$\mathrm{M}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

be the matrix for reflection through the x_1-axis.

Then following the above sequence of maps the matrix of the reflection is

$$\mathrm{R}_\theta \mathrm{M}_1 \mathrm{R}_\theta^{-1}\,.$$

Another Way to Reflect About a Line in $\mathbb{R}^2$

Challenge Problem: Show that the reflection through a line $\ell = \mathrm{Span}\left\{\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}\right\}$ is given by

$$\mathrm{Ref}_\ell(\mathbf{x}) = \begin{bmatrix}\cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta\end{bmatrix}\mathbf{x}\,,$$

by computing the product $\mathrm{R}_\theta \mathrm{M}_1 \mathrm{R}_\theta^{-1}$. Compare with the formula obtained by computing $\mathrm{Ref}_\ell(\mathbf{x}) = (2\mathrm{proj}_\mathbf{v} - \mathrm{I}_2)\mathbf{x}$.

In terms of complex numbers $z = x + iy$, one can rewrite this using the conjugation map $z \mapsto \bar{z} = x - yi$. Indeed, the reflection through $\ell = \{se^{i\theta} : s \in \mathbb{R}\}$ is given by

$$z \mapsto e^{i\theta}\overline{(e^{-i\theta}z)} = e^{2i\theta}\bar{z}\,.$$

Equivalent conditions for Invertibility

For a linear transformation $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ of $\mathbb{R}^n$:

- The map $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ is surjective (onto) if and only if the columns of A span $\mathbb{R}^n$, if and only if there is a pivot in each row of A, if and only if here is a solution to the system $\mathrm{A}\mathbf{x} = \mathbf{b}$ for any $\mathbf{b}$ in $\mathbb{R}^n$.
- The map $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ is injective (one-to-one) if and only if $\ker T$ is trivial, if and only if the columns of A are linearly independent, if and only if there is a pivot position in each column of A.

Since A is square, there are pivots in each column if and only if there are pivots in each row, and so A is either both injective and surjective, or it is neither.
And the map $T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ is invertible if and only if it is bijective (both one-to-one and onto).

The Invertible Matrix Theorem

Theorem

Let $\mathrm{A} \in \mathbb{R}^{n\times n}$ *The following statements are all logically equivalent.*

a. *The matrix* A *is invertible.*

b. $\mathrm{RREF(A)} = \mathrm{I}_n$.

c. A *has* n *pivot positions.*

d. $\ker \mathrm{A} = \{\mathbf{0}\}$.

e. *The columns of* A *are linearly independent.*

f. *The linear transformation* $\mathbf{x} \mapsto \mathrm{A}\mathbf{x}$ *is injective.*

g. *The equation* $\mathrm{A}\mathbf{x} = \mathbf{b}$ *has a solution for each* $\mathbf{b} \in \mathbb{R}^n$.

h. *The columns of* A *span* $\mathbb{R}^n$.

i. *The linear transformation* $\mathbf{x} \mapsto \mathrm{A}\mathbf{x}$ *maps* $\mathbb{R}^n$ *onto* $\mathbb{R}^n$.

j. *There exists* $\mathrm{B} \in \mathbb{R}^{n\times n}$ *such that* $\mathrm{BA} = \mathrm{I}_n$.

k. *There exists* $\mathrm{C} \in \mathbb{R}^{n\times n}$ *such that* $\mathrm{AC} = \mathrm{I}_n$.

l. A^{t} *is invertible.*

Proof of the Invertible matrix theorem

Proof.

Many of the details have already been shown; it remains to string together the appropriate circles of logical implication. We defer the details of the proof to the next class. You can also read a proof in the text, section 2.3. □

Left and Right Inverses: the Rectangular Case

If $\mathrm{A} \in \mathbb{R}^{m\times n}$, there won't be a total inverse matrix, but there may be either a left or right inverse, in the sense described above for functions.

Challenge Problem: Find conditions on an $m \times n$ matrix A for the existence of a left or right inverse. How do you find such an inverse? You'll want to look carefully at the algorithm for square matrices, and at the invertible matrix theorem, and figure out what must be different for the non-square cases.

Exam Details

- Midterm exam 1 is in **room 124 in the Hasbrouck Lab Addition**, tomorrow night, from **7-9 pm**.
- You may bring one letter size ($8.5'' \times 11''$) sheet of notes (both sides). No additional notes, no textbooks, and no calculators are allowed.
- The exam covers 1.1, 1.2, 1.3, 1.4, 1.5, 1.7, 1.8, 1.9, and 2.1 in the text.

In particular, you should be comfortable with the following:

- Writing linear systems using equations or matrices, using vector arithmetic (addition and scaling), and using the language of linear combinations.
- Performing row reduction, interpreting the reduced row echelon form, and solving systems. Know what it means to have a pivot in each row, or a pivot in each column.
- Knowing how to find dependence relations, solving homogeneous and inhomogeneous systems, and understanding linear independence (both the definition and how to test if a collection of vectors is linearly independent).
- Computing matrix-vector products, multiplying matrices, and representing linear maps by matrices.
- Generally, knowing all of the definitions from the relevant sections, and being able to explain them in your own words.