Matrix Calculations: Linear maps, bases, and matrix multiplication

A. Kissinger

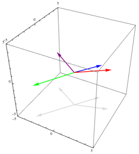
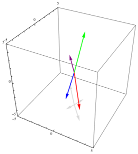
Institute for Computing and Information Sciences
Radboud University Nijmegen

Version: autumn 2018

Outline

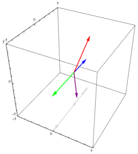
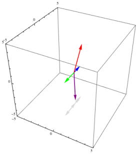
From last time

- Linear maps describe *transformations in space*, such as rotation:

 $\mapsto$ 

$$\mathrm{rx}\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y\cos\theta - z\sin\theta \\ y\sin\theta + z\cos\theta \end{pmatrix}$$

- reflection and scaling:

 $\mapsto$ 

$$\mathrm{sy}\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ (1/2)y \\ z \end{pmatrix}$$

From last time

- Linear maps can be represented as a matrix, using matrix multiplication:

$$f(\boldsymbol{v}) = \boldsymbol{A} \cdot \boldsymbol{v}$$

- For example, then linear map:

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y\cos\theta - z\sin\theta \\ y\sin\theta + z\cos\theta \end{pmatrix}$$

can be represented as:

$$f(\underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\boldsymbol{v}}) = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}}_{\boldsymbol{A}} \cdot \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\boldsymbol{v}}$$

Matrix multiplication

- Consider linear maps g, f represented by matrices $\boldsymbol{A}$, $\boldsymbol{B}$:

$$g(\boldsymbol{v}) = \boldsymbol{A} \cdot \boldsymbol{v} \qquad\qquad f(\boldsymbol{w}) = \boldsymbol{B} \cdot \boldsymbol{w}$$

- Can we find a matrix $\boldsymbol{C}$ that represents their composition?

$$g(f(\boldsymbol{v})) = \boldsymbol{C} \cdot \boldsymbol{v}$$

- Let's try:

$$g(f(\boldsymbol{v})) = g(\boldsymbol{B} \cdot \boldsymbol{v}) = \boldsymbol{A} \cdot (\boldsymbol{B} \cdot \boldsymbol{v}) \overset{(*)}{=} (\boldsymbol{A} \cdot \boldsymbol{B}) \cdot \boldsymbol{v}$$

(where step $(*)$ is currently 'wishful thinking')

- Great! Let $\boldsymbol{C} := \boldsymbol{A} \cdot \boldsymbol{B}$.
- But we don't know what "$\cdot$" means for two matrices yet...

Matrix multiplication

- Solution: generalise from $\boldsymbol{A} \cdot \boldsymbol{v}$
- A vector is a matrix with one column:

  The number in the i-th row and the first column of $\boldsymbol{A} \cdot \boldsymbol{v}$ is the dot product of the i-th row of $\boldsymbol{A}$ with the first column of $\boldsymbol{v}$.
- So for matrices $\boldsymbol{A}, \boldsymbol{B}$:

  The number in the i-th row and the j-th column of $\boldsymbol{A} \cdot \boldsymbol{B}$ is the dot product of the i-th row of $\boldsymbol{A}$ with the j-th column of $\boldsymbol{B}$.

Matrix multiplication

For $\mathbf{A}$ an $m \times n$ matrix, $\mathbf{B}$ an $n \times p$ matrix:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{C}$$

is an $m \times p$ matrix.

$$\begin{pmatrix} \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} \cdots & b_{j1} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & b_{jn} & \cdots \end{pmatrix} = \begin{pmatrix} \ddots & \vdots & \iddots \\ \cdots & c_{ij} & \cdots \\ \iddots & \vdots & \ddots \end{pmatrix}$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Special case: vectors

For $\boldsymbol{A}$ an $m \times n$ matrix, $\boldsymbol{B}$ an $n \times 1$ matrix:

$$\boldsymbol{A} \cdot \boldsymbol{b} = \boldsymbol{c}$$

is an $m \times 1$ matrix.

$$\begin{pmatrix} \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} \vdots \\ c_{i1} \\ \vdots \end{pmatrix}$$

$$c_{i1} = \sum_{k=1}^{n} a_{ik} b_{k1}$$

Matrix composition

Theorem

Matrix composition is associative:

$$(\boldsymbol{A} \cdot \boldsymbol{B}) \cdot \boldsymbol{C} = \boldsymbol{A} \cdot (\boldsymbol{B} \cdot \boldsymbol{C})$$

Proof. Let $\boldsymbol{X} := \boldsymbol{A} \cdot \boldsymbol{B}$. This is a matrix with entries:

$$x_{ip} = \sum_k a_{ik} b_{kp}$$

Then, the matrix entries of $\boldsymbol{X} \cdot \boldsymbol{C}$ are:

$$\sum_p x_{ip} c_{pj} = \sum_p \left(\sum_k a_{ik} b_{kp} \right) c_{pj} = \sum_{kp} a_{ik} b_{kp} c_{pj}$$

(because sums can always be pulled outside, and combined)

Associativity of matrix composition

Proof (cont'd). Now, let $\boldsymbol{Y} := \boldsymbol{B} \cdot \boldsymbol{C}$. This has matrix entries:

$$y_{kj} = \sum_p b_{kp} c_{pj}$$

Then, the matrix entries of $\boldsymbol{A} \cdot \boldsymbol{Y}$ are:

$$\sum_k a_{ik} y_{kj} = \sum_k a_{ik} \left(\sum_p b_{kp} c_{pj} \right) = \sum_{kp} a_{ik} b_{kp} c_{pj}$$

...which is the same as before! So:

$$(\boldsymbol{A} \cdot \boldsymbol{B}) \cdot \boldsymbol{C} = \boldsymbol{X} \cdot \boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{Y} = \boldsymbol{A} \cdot (\boldsymbol{B} \cdot \boldsymbol{C})$$

☺

So we can drop those pesky parentheses:

$$\boldsymbol{A} \cdot \boldsymbol{B} \cdot \boldsymbol{C} := (\boldsymbol{A} \cdot \boldsymbol{B}) \cdot \boldsymbol{C} = \boldsymbol{A} \cdot (\boldsymbol{B} \cdot \boldsymbol{C})$$

Matrix product and composition

Corollary

The composition of linear maps is given by matrix product.

Proof. Let $g(\boldsymbol{w}) = \boldsymbol{A} \cdot \boldsymbol{w}$ and $f(\boldsymbol{v}) = \boldsymbol{B} \cdot \boldsymbol{v}$. Then:

$$g(f(\boldsymbol{v})) = g(\boldsymbol{B} \cdot \boldsymbol{v}) = \boldsymbol{A} \cdot \boldsymbol{B} \cdot \boldsymbol{v}$$

☺

No wishful thinking necessary!

Example 1

Consider the following two linear maps, and their associated matrices:

$$\mathbb{R}^3 \xrightarrow{f} \mathbb{R}^2 \qquad \mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2$$

$$f((x_1, x_2, x_3)) = (x_1 - x_2,\ x_2 + x_3) \qquad g((y_1, y_2)) = (2y_1 - y_2, 3y_2)$$

$$\boldsymbol{M}_f = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad \boldsymbol{M}_g = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$$

We can compute the composition directly:

$$\begin{aligned}(g \circ f)((x_1, x_2, x_3)) &= g\big(f((x_1, x_2, x_3))\big) \\ &= g((x_1 - x_2,\ x_2 + x_3)) \\ &= (\,2(x_1 - x_2) - (x_2 + x_3),\ 3(x_2 + x_3)\,) \\ &= (\,2x_1 - 3x_2 - x_3,\ 3x_2 + 3x_3\,)\end{aligned}$$

So:

$$\boldsymbol{M}_{g \circ f} = \begin{pmatrix} 2 & -3 & -1 \\ 0 & 3 & 3 \end{pmatrix}$$

...which is just the product of the matrices: $\boldsymbol{M}_{g \circ f} = \boldsymbol{M}_g \cdot \boldsymbol{M}_f$

Note: matrix composition is not commutative

In general, $\boldsymbol{A} \cdot \boldsymbol{B} \neq \boldsymbol{B} \cdot \boldsymbol{A}$

For instance: Take $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then:

$$\begin{aligned} \boldsymbol{A} \cdot \boldsymbol{B} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 0 + 0 \cdot -1 & 1 \cdot 1 + 0 \cdot 0 \\ 0 \cdot 0 + -1 \cdot -1 & 0 \cdot 1 + -1 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \boldsymbol{B} \cdot \boldsymbol{A} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \cdot 1 + 1 \cdot 0 & 0 \cdot 0 + 1 \cdot -1 \\ -1 \cdot 1 + 0 \cdot 0 & -1 \cdot 0 + 0 \cdot -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \end{aligned}$$

But it is...

...associative, as we've already seen:

$$\boldsymbol{A} \cdot \boldsymbol{B} \cdot \boldsymbol{C} := (\boldsymbol{A} \cdot \boldsymbol{B}) \cdot \boldsymbol{C} = \boldsymbol{A} \cdot (\boldsymbol{B} \cdot \boldsymbol{C})$$

It also has a *unit* given by the *identity matrix* $\boldsymbol{I}$:

$$\boldsymbol{A} \cdot \boldsymbol{I} = \boldsymbol{I} \cdot \boldsymbol{A} = \boldsymbol{A}$$

where:

$$\boldsymbol{I} := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Solving equations the old fashioned way...

- We now know that systems of equations look like this:

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$$

- The goal is to solve for $\boldsymbol{x}$, in terms of $\boldsymbol{A}$ and $\boldsymbol{b}$.
- Here comes some more wishful thinking:

$$\boldsymbol{x} = \frac{1}{\boldsymbol{A}} \cdot \boldsymbol{b}$$

- Well, we can't really *divide* by a matrix, but if we are lucky, we can find another matrix called $\boldsymbol{A}^{-1}$ which acts like $\frac{1}{\boldsymbol{A}}$.

Inverse

Definition

The *inverse* of a matrix $\boldsymbol{A}$ is another matrix $\boldsymbol{A}^{-1}$ such that:

$$\boldsymbol{A}^{-1} \cdot \boldsymbol{A} = \boldsymbol{A} \cdot \boldsymbol{A}^{-1} = \boldsymbol{I}$$

- Not all matrices have inverses, but when they do, we are happy, because:

$$\begin{aligned} \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b} \quad &\Longrightarrow \quad \boldsymbol{A}^{-1} \cdot \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b} \\ &\Longrightarrow \quad \boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b} \end{aligned}$$

- So, how do we compute the inverse of a matrix?

Remember me?

Gaussian elimination as matrix multiplication

- Each step of Gaussian elimination can be represented by a matrix multiplication:

$$\boldsymbol{A} \Rightarrow \boldsymbol{A}' \qquad\qquad \boldsymbol{A}' := \boldsymbol{G} \cdot \boldsymbol{A}$$

- For instance, multiplying the i-th row by c is given by:

$$\boldsymbol{G}_{(R_i := cR_i)} \cdot \boldsymbol{A}$$

where $\boldsymbol{G}_{(R_i := cR_i)}$ is just like the identity matrix, but $g_{ii} = c$.

- Exercise. What are the other Gaussian elimination matrices?

$$\boldsymbol{G}_{(R_i \leftrightarrow R_j)} \qquad\qquad \boldsymbol{G}_{(R_i := R_i + cR_j)}$$

Reduction to Echelon form

- The idea: treat $\boldsymbol{A}$ as a coefficient matrix, and compute its reduced Echelon form
- If the Echelon form of $\boldsymbol{A}$ has n pivots, then its reduced Echelon form is the identity matrix:

$$\boldsymbol{A} \Rightarrow \boldsymbol{A}_1 \Rightarrow \boldsymbol{A}_2 \Rightarrow \cdots \Rightarrow \boldsymbol{A}_p = \boldsymbol{I}$$

- Now, we can use our Gauss matrices to remember what we did:

$$\begin{aligned}
\boldsymbol{A}_1 &:= \boldsymbol{G}_1 \cdot \boldsymbol{A} \\
\boldsymbol{A}_2 &:= \boldsymbol{G}_2 \cdot \boldsymbol{G}_1 \cdot \boldsymbol{A} \\
&\cdots \\
\boldsymbol{A}_p &:= \boldsymbol{G}_p \cdots \boldsymbol{G}_1 \cdot \boldsymbol{A} = \boldsymbol{I}
\end{aligned}$$

Computing the inverse

- A ha!

$$\mathbf{G}_p \cdots \mathbf{G}_1 \cdot \mathbf{A} = \mathbf{I} \qquad \Longrightarrow \qquad \mathbf{A}^{-1} = \mathbf{G}_p \cdots \mathbf{G}_1$$

- So all we have to do is construct p different matrices and multiply them all together!
- Since I already have plans for this afternoon, lets take a shortcut.

Computing the inverse

- Since Gaussian elimination is just multiplying by a certain matrix on the left...

$$\boldsymbol{A} \Rightarrow \boldsymbol{G} \cdot \boldsymbol{A}$$

- ...doing Gaussian elimination (for $\boldsymbol{A}$) on an augmented matrix applies $\boldsymbol{G}$ to both parts:

$$(\boldsymbol{A}|\boldsymbol{B}) \Rightarrow (\boldsymbol{G} \cdot \boldsymbol{A} \mid \boldsymbol{G} \cdot \boldsymbol{B})$$

- So, if $\boldsymbol{G} = \boldsymbol{A}^{-1}$:

$$(\boldsymbol{A}|\boldsymbol{B}) \Rightarrow (\boldsymbol{A}^{-1} \cdot \boldsymbol{A} \mid \boldsymbol{A}^{-1} \cdot \boldsymbol{B}) = (\boldsymbol{I}|\ \boldsymbol{A}^{-1} \cdot \boldsymbol{B})$$

Computing the inverse

- We already (secretly) used this trick to solve:

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b} \qquad \Longrightarrow \qquad \boldsymbol{x} = \boldsymbol{A}^{-1} \cdot \boldsymbol{b}$$

- Here, we are only interested in the vector $\boldsymbol{A}^{-1} \cdot \boldsymbol{b}$
- Which is exactly what Gaussian elimination on the augmented matrix gives us:

$$(\boldsymbol{A}|\boldsymbol{b}) \Rightarrow (\boldsymbol{I}|\ \boldsymbol{A}^{-1} \cdot \boldsymbol{b})$$

- To get the entire matrix, we just need to choose something clever to the right of the line
- Like this:

$$(\boldsymbol{A}|\boldsymbol{I}) \Rightarrow (\boldsymbol{I}|\ \boldsymbol{A}^{-1} \cdot \boldsymbol{I}) = (\boldsymbol{I}|\ \boldsymbol{A}^{-1})$$

Computing the inverse: example

For example, we compute the inverse of:

$$\boldsymbol{A} := \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

as follows:

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

So:

$$\boldsymbol{A}^{-1} := \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

Computing the inverse: non-example

Unlike transpose, not every matrix has an inverse.
For example, if we try to compute the inverse for:

$$\boldsymbol{B} := \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

we have:

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{array}\right)$$

We don't have enough pivots to continue reducing. So $\boldsymbol{B}$ does not have an inverse.

When does a matrix have an inverse?

Theorem (Existence of inverses)

An $n \times n$ matrix has an inverse (or: is invertible) if and only if it has n pivots in its echelon form.

Next time, we will introduce another criterion for a matrix to be invertible, using determinants.

Uniqueness of the inverse

Note

Matrix multiplication is not commutative, so it could (*a priori*) be the case that:

- $\boldsymbol{A}$ has a right inverse: a $\boldsymbol{B}$ such that $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{I}$ and
- $\boldsymbol{A}$ has a (different) left inverse: a $\boldsymbol{C}$ such that $\boldsymbol{C} \cdot \boldsymbol{A} = \boldsymbol{I}$.

However, this doesn't happen.

Uniqueness of the inverse

Theorem

If a matrix $\mathbf{A}$ has a left inverse and a right inverse, then they are equal. If $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}$ and $\mathbf{C} \cdot \mathbf{A} = \mathbf{I}$, then $\mathbf{B} = \mathbf{C}$.

Proof. Multiply both sides of the first equation by $\mathbf{C}$:

$$\mathbf{C} \cdot \mathbf{A} \cdot \mathbf{B} = \mathbf{C} \cdot \mathbf{I} \qquad \Longrightarrow \qquad \mathbf{B} = \mathbf{C}$$

Corollary

If a matrix $\mathbf{A}$ has an inverse, it is unique.

Explicitly computing the inverse, part I

- Suppose we wish to find $\mathbf{A}^{-1}$ for $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
- We need to find x, y, u, v with:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Multiplying the matrices on the LHS:

$$\begin{pmatrix} ax + bu & ay + bv \\ cx + du & cy + dv \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- ...gives a system of 4 equations:

$$\begin{cases} ax + bu &= 1 \\ cx + du &= 0 \\ ay + bv &= 0 \\ cy + dv &= 1 \end{cases}$$

Computing the inverse: the 2×2 case, part II

- Splitting this into two systems:

$$\begin{cases} ax + bu = 1 \\ cx + du = 0 \end{cases} \quad \text{and} \quad \begin{cases} ay + bv = 0 \\ cy + dv = 1 \end{cases}$$

- Solving the first system for (u, x) and the second system for (v, y) gives:

$$u = \frac{-c}{ad-bc} \quad x = \frac{d}{ad-bc} \quad \text{and} \quad v = \frac{a}{ad-bc} \quad y = \frac{-b}{ad-bc}$$

(assuming $ad - bc \neq 0$). Then:

$$\boldsymbol{A}^{-1} = \begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}$$

- Conclusion: $\boldsymbol{A}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ ✓

learn this formula by heart

Computing the inverse: the 2×2 case, part III

Summarizing:

Theorem (Existence of an inverse of a 2×2 matrix)

A 2×2 matrix

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has an inverse (or: is invertible) if and only if $ad - bc \neq 0$, in which case its inverse is

$$\boldsymbol{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Example

- Let $\boldsymbol{P} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$, so $a = \frac{8}{10}, b = \frac{1}{10}, c = \frac{2}{10}, d = \frac{9}{10}$
- $ad - bc = \frac{72}{100} - \frac{2}{100} = \frac{70}{100} = \frac{7}{10} \neq 0$ so the inverse exists!
- Thus:

$$\begin{aligned} \boldsymbol{P}^{-1} &= \tfrac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\ &= \tfrac{10}{7} \begin{pmatrix} 0.9 & -0.1 \\ -0.2 & 0.8 \end{pmatrix} \end{aligned}$$

- Then indeed:

$$\tfrac{10}{7} \begin{pmatrix} 0.9 & -0.1 \\ -0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix} = \tfrac{10}{7} \begin{pmatrix} 0.7 & 0 \\ 0 & 0.7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- You could try to do this for bigger matrices, but it's very complicated. $\Longrightarrow$ Gauss elimination is way easier!