Matrix Calculations: Inverse and Basis Transformation

J. van de Wetering (and A. Kissinger)

Institute for Computing and Information Sciences
Radboud University Nijmegen

spring 2017

Outline

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:

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

Computing the inverse

- A ha!

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

- 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:

Theorem

For $\boldsymbol{C}$ *a matrix and* $(\boldsymbol{A}|\boldsymbol{B})$ *an augmented matrix:*

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

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.

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

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

$$\boldsymbol{C} \cdot \boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{C} \cdot \boldsymbol{I} \quad \Longrightarrow \quad \boldsymbol{B} = \boldsymbol{C}$$

☺

Corollary

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

Explicitly computing the inverse, part I

- Suppose we wish to find $\boldsymbol{A}^{-1}$ for $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
- 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:

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

- ...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 $bc - ad \neq 0$). 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:

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

Determinants

What a determinant does

For a square matrix $\boldsymbol{A}$, the determinant $\det(\boldsymbol{A})$ is a number (in $\mathbb{R}$)

It satisfies:

$$\begin{aligned} \det(\boldsymbol{A}) = 0 &\iff \boldsymbol{A} \text{ is not invertible} \\ &\iff \boldsymbol{A}^{-1} \text{ does not exist} \\ &\iff \boldsymbol{A} \text{ has} < n \text{ pivots in its echolon form} \end{aligned}$$

Determinants have useful properties, but calculating determinants involves some work.

Determinant of a 2×2 matrix

- Assume $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
- Recall that the inverse $\boldsymbol{A}^{-1}$ exists if and only if $ad - bc \neq 0$, and in that case is:

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

- In this 2×2-case we define:

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

- Thus, indeed: $\det(\boldsymbol{A}) = 0 \iff \boldsymbol{A}^{-1}$ does not exist.

Determinant of a 2×2 matrix: example

- Recall the political transisition matrix

$$\boldsymbol{P} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 8 & 1 \\ 2 & 9 \end{pmatrix}$$

- Then:

$$\begin{aligned} \det(\boldsymbol{P}) &= \frac{8}{10} \cdot \frac{9}{10} - \frac{1}{10} \cdot \frac{2}{10} \\ &= \frac{72}{100} - \frac{2}{100} \\ &= \frac{70}{100} = \frac{7}{10} \end{aligned}$$

- We have already seen that $\boldsymbol{P}^{-1}$ exists, so the determinant must be non-zero.

Determinant of a 3×3 matrix

- Assume $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$
- Then one defines:

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = +a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \cdot \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \cdot \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

- Methodology:
  - take entries a_{i1} from first column, with alternating signs (+, -)
  - take determinant from square submatrix obtained by deleting the first column and the i-th row

Determinant of a 3×3 matrix, example

$$\begin{vmatrix} 1 & 2 & -1 \\ 5 & 3 & 4 \\ -2 & 0 & 1 \end{vmatrix} = 1\begin{vmatrix} 3 & 4 \\ 0 & 1 \end{vmatrix} - 5\begin{vmatrix} 2 & -1 \\ 0 & 1 \end{vmatrix} + -2\begin{vmatrix} 2 & -1 \\ 3 & 4 \end{vmatrix}$$

$$= \big(3-0\big) - 5\big(2-0\big) - 2\big(8+3\big)$$

$$= 3 - 10 - 22$$

$$= -29$$

The general, $n \times n$ case

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = +a_{11} \cdot \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix} - a_{21} \cdot \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ a_{32} & \cdots & a_{3n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

$$+ a_{31} \begin{vmatrix} \ldots \\ \ldots \\ \ldots \end{vmatrix} \quad \cdots \quad \pm a_{n1} \begin{vmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{(n-1)2} & \ldots & a_{(n-1)n} \end{vmatrix}$$

(where the last sign $\pm$ is $+$ if n is odd and - if n is even)

Then, each of the smaller determinants is computed recursively.

(A lot of work! But there are smarter ways...)

Some properties of determinants

Theorem

For $\boldsymbol{A}$ and $\boldsymbol{B}$ two $n \times n$ matrices,

$$\det(\boldsymbol{A} \cdot \boldsymbol{B}) = \det(\boldsymbol{A}) \cdot \det(\boldsymbol{B}).$$

The following are corollaries of the Theorem:

- $\det(\boldsymbol{A} \cdot \boldsymbol{B}) = \det(\boldsymbol{B} \cdot \boldsymbol{A})$.
- If $\boldsymbol{A}$ has an inverse, then $\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$.
- $\det(\boldsymbol{A}^k) = (\det(\boldsymbol{A}))^k$, for any $k \in \mathbb{N}$.

Proofs of the first two:

- $\det(\boldsymbol{A} \cdot \boldsymbol{B}) = \det(\boldsymbol{A}) \cdot \det(\boldsymbol{B}) = \det(\boldsymbol{B}) \cdot \det(\boldsymbol{A}) = \det(\boldsymbol{B} \cdot \boldsymbol{A})$. (Note that $\det(\boldsymbol{A})$ and $\det(\boldsymbol{B})$ are simply numbers).
- If $\boldsymbol{A}$ has an inverse $\boldsymbol{A}^{-1}$ then
  $\det(\boldsymbol{A}) \cdot \det(\boldsymbol{A}^{-1}) = \det(\boldsymbol{A} \cdot \boldsymbol{A}^{-1}) = \det(\boldsymbol{I}) = 1$, so
  $\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$.

Applications

- Determinants detect when a matrix is invertible
- Though we showed an inefficient way to compute determinants, there is an efficient algorithm using, you guessed it...Gaussian elimination!
- Solutions to non-homogeneous systems can be expressed directly in terms of determinants using *Cramer's rule* (wiki it!)
- Most importantly: determinants will be used to calculate *eigenvalues* in the next lecture

Vectors in a basis

Recall: a basis for a vector space V is a set of vectors $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ in V such that:

1. They **uniquely** *span* V, i.e. for all $\boldsymbol{v} \in V$, there exist **unique** a_i such that:

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n$$

Because of this, we use a special notation for this linear combination:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}} := a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n$$

Same vector, different outfits

The *same vector* can look different, depending on the choice of basis:

$$\begin{pmatrix} 100 \cdot (a+b) \\ b \end{pmatrix}_{\mathcal{S}} = \begin{pmatrix} a \\ b \end{pmatrix}_{\mathcal{B}}$$

Examples:

$$\begin{pmatrix} 100 \\ 0 \end{pmatrix}_{\mathcal{S}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{B}} \qquad \begin{pmatrix} 300 \\ 1 \end{pmatrix}_{\mathcal{S}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{\mathcal{B}}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{S}} = \begin{pmatrix} \frac{1}{100} \\ 0 \end{pmatrix}_{\mathcal{B}} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{S}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}_{\mathcal{B}}$$

Transforming bases, part I

- **Problem:** given a vector written in $\mathcal{B} = \{(100, 0), (100, 1)\}$, how can we write it in the standard basis? Just use the definition:

$$\begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{B}} = x \cdot \begin{pmatrix} 100 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 100 \\ 1 \end{pmatrix} = \begin{pmatrix} 100x + 100y \\ y \end{pmatrix}_{\mathcal{S}}$$

- Or, as matrix multiplication:

$$\underbrace{\begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix}}_{\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}}} \cdot \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\text{in basis } \mathcal{B}} = \underbrace{\begin{pmatrix} 100x + 100y \\ y \end{pmatrix}}_{\text{in basis } \mathcal{S}}$$

- Let $\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}}$ be the matrix whose *columns* are the basis vectors $\mathcal{B}$. Then $\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}}$ *transforms* a vector written in $\mathcal{B}$ into a vector written in $\mathcal{S}$.

Transforming bases, part II

- How do we transform back? Need $\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}}$ which undoes the matrix $\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}}$.
- Solution: use the inverse! $\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}} := (\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}})^{-1}$
- Example:

$$(\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}})^{-1} = \begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{100} & -1 \\ 0 & 1 \end{pmatrix}$$

- ...which indeed gives:

$$\begin{pmatrix} \frac{1}{100} & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{a-100b}{100} \\ b \end{pmatrix}$$

Transforming bases, part IV

- How about two non-standard bases?

$$\mathcal{B} = \left\{ \begin{pmatrix} 100 \\ 0 \end{pmatrix}, \begin{pmatrix} 100 \\ 1 \end{pmatrix} \right\} \qquad \mathcal{C} = \left\{ \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

- **Problem**: translate a vector from $\begin{pmatrix} a \\ b \end{pmatrix}_{\mathcal{B}}$ to $\begin{pmatrix} a' \\ b' \end{pmatrix}_{\mathcal{C}}$
- **Solution**: do this in two steps:

$$\underbrace{\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}} \cdot \boldsymbol{v}}_{\text{first translate from } \mathcal{B} \text{ to } \mathcal{S}\ldots}$$

$$\underbrace{\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{C}} \cdot \boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}} \cdot \boldsymbol{v}}_{\ldots\text{then translate from } \mathcal{S} \text{ to } \mathcal{C}} \quad = \quad \left(\boldsymbol{T}_{\mathcal{C}\Rightarrow\mathcal{S}}\right)^{-1} \cdot \boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}} \cdot \boldsymbol{v}$$

Transforming bases, example

- For bases:

$$\mathcal{B} = \left\{ \begin{pmatrix} 100 \\ 0 \end{pmatrix}, \begin{pmatrix} 100 \\ 1 \end{pmatrix} \right\} \qquad \mathcal{C} = \left\{ \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

- ...we need to find a' and b' such that

$$\begin{pmatrix} a' \\ b' \end{pmatrix}_{\mathcal{C}} = \begin{pmatrix} a \\ b \end{pmatrix}_{\mathcal{B}}$$

- Translating both sides to the standard basis gives:

$$\begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}$$

- This we can solve using the matrix-inverse:

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}$$

Transforming bases, example

For:

$$\underbrace{\begin{pmatrix} a' \\ b' \end{pmatrix}}_{\text{in basis } \mathcal{C}} = \underbrace{\begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix}^{-1}}_{\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{C}}} \cdot \underbrace{\begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix}}_{\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}}} \cdot \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{\text{in basis } \mathcal{B}}$$

we compute

$$\begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} \end{pmatrix} \cdot \begin{pmatrix} 100 & 100 \\ 0 & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -200 & -199 \\ 200 & 201 \end{pmatrix}$$

which gives:

$$\underbrace{\begin{pmatrix} a' \\ b' \end{pmatrix}}_{\text{in basis } \mathcal{C}} = \underbrace{\frac{1}{4} \begin{pmatrix} -200 & -199 \\ 200 & 201 \end{pmatrix}}_{\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{C}}} \cdot \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{\text{in basis } \mathcal{B}}$$

Basis transformation theorem

Theorem

Let $\mathcal{S}$ be the standard basis for $\mathbb{R}^n$ and let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be other bases.

1. *Then there is an invertible $n \times n$ basis transformation matrix $\mathbf{T}_{\mathcal{B}\Rightarrow\mathcal{C}}$ such that:*

$$\begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix} = \mathbf{T}_{\mathcal{B}\Rightarrow\mathcal{C}} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \textit{with} \quad \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix}_{\mathcal{C}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}}$$

2. *$\mathbf{T}_{\mathcal{B}\Rightarrow\mathcal{S}}$ is the matrix which has the vectors in $\mathcal{B}$ as columns, and*

$$\mathbf{T}_{\mathcal{B}\Rightarrow\mathcal{C}} := (\mathbf{T}_{\mathcal{C}\Rightarrow\mathcal{S}})^{-1} \cdot \mathbf{T}_{\mathcal{B}\Rightarrow\mathcal{S}}$$

3. $\mathbf{T}_{\mathcal{C}\Rightarrow\mathcal{B}} = (\mathbf{T}_{\mathcal{B}\Rightarrow\mathcal{C}})^{-1}$

Matrices in other bases

- Since *vectors* can be written with respect to different bases, so too can *matrices*.
- For example, let g be the linear map defined by:

$$g\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{S}} \right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{S}} \qquad g\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{S}} \right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{S}}$$

- Then, naturally, we would represent g using the matrix:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_{\mathcal{S}}$$

- Because indeed:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

On the other hand...

- Lets look at what g does to another basis:

$$\mathcal{B} = \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\}$$

- First $(1,1) \in \mathcal{B}$:

$$g(\ \begin{pmatrix}1\\0\end{pmatrix}_{\mathcal{B}}\ ) = g(\ \begin{pmatrix}1\\1\end{pmatrix}\ ) = g(\ \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\ ) = \dots$$

- Then, by linearity:

$$\dots = g(\ \begin{pmatrix}1\\0\end{pmatrix}\ ) + g(\ \begin{pmatrix}0\\1\end{pmatrix}\ ) = \begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}_{\mathcal{B}}$$

On the other hand...

$$\mathcal{B} = \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\}$$

- Similarly $(1,-1) \in \mathcal{B}$:

$$g(\begin{pmatrix}0\\1\end{pmatrix}_{\mathcal{B}}) = g(\begin{pmatrix}1\\-1\end{pmatrix}) = g(\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix}) = \dots$$

- Then, by linearity:

$$\dots = g(\begin{pmatrix}1\\0\end{pmatrix}) - g(\begin{pmatrix}0\\1\end{pmatrix}) = \begin{pmatrix}0\\1\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix} = -\begin{pmatrix}0\\1\end{pmatrix}_{\mathcal{B}}$$

A new matrix

- From this:

$$g\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{B}} \right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{B}} \qquad g\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{B}} \right) = - \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{B}}$$

- It follows that we should instead use *this* matrix to represent g:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mathcal{B}}$$

- Because indeed:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = - \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

A new matrix

- So on different bases, g acts in a totally different way!

$$g(\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{S}}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{S}} \qquad g(\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{S}}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{S}}$$

$$g(\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{B}}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{B}} \qquad g(\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{B}}) = -\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{B}}$$

- ...and hence gets a totally different matrix:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_{\mathcal{S}} \quad \text{vs.} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{\mathcal{B}}$$

Transforming bases, part II

Theorem

Assume again we have two bases $\mathcal{B}, \mathcal{C}$ for $\mathbb{R}^n$.

If a linear map $f : \mathbb{R}^n \to \mathbb{R}^n$ has matrix $\boldsymbol{A}$ w.r.t. to basis $\mathcal{B}$, then, w.r.t. to basis $\mathcal{C}$, f has matrix $\boldsymbol{A}'$:

$$\boldsymbol{A}' = \boldsymbol{T}_{\mathcal{B} \Rightarrow \mathcal{C}} \cdot \boldsymbol{A} \cdot \boldsymbol{T}_{\mathcal{C} \Rightarrow \mathcal{B}}$$

Thus, via $\boldsymbol{T}_{\mathcal{B} \Rightarrow \mathcal{C}}$ and $\boldsymbol{T}_{\mathcal{C} \Rightarrow \mathcal{B}}$ one tranforms $\mathcal{B}$-matrices into $\mathcal{C}$-matrices. In particular, a matrix can be translated from the standard basis to basis $\mathcal{B}$ via:

$$\boldsymbol{A}' = \boldsymbol{T}_{\mathcal{S} \Rightarrow \mathcal{B}} \cdot \boldsymbol{A} \cdot \boldsymbol{T}_{\mathcal{B} \Rightarrow \mathcal{S}}$$

Example basis transformation, part I

- Consider the standard basis $\mathcal{S} = \{(1,0),(0,1)\}$ for $\mathbb{R}^2$, and as alternative basis $\mathcal{B} = \{(-1,1),(0,2)\}$
- Let the linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$, w.r.t. the standard basis $\mathcal{S}$, be given by the matrix:

$$\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$$

- What is the representation $\boldsymbol{A}'$ of f w.r.t. basis $\mathcal{B}$?
- Since $\mathcal{S}$ is the standard basis, $\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}} = \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix}$ contains the $\mathcal{B}$-vectors as its columns