Matrix Calculations: Inverse and Basis Transformation

A. Kissinger (and H. Geuvers)

Institute for Computing and Information Sciences – Intelligent Systems
Radboud University Nijmegen

Version: spring 2015

Outline

Recall: Inverse matrix

Definition

Let $\boldsymbol{A}$ be a $n \times n$ ("square") matrix.

This $\boldsymbol{A}$ has an inverse if there is an $n \times n$ matrix $\boldsymbol{A}^{-1}$ with:

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

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.

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.

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:

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

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

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

Applying the general formula to the swingers

- Recall $\boldsymbol{P} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$, 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 deteminant $\det(\boldsymbol{A})$ 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:

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

- 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

$$
\begin{aligned}
\begin{vmatrix} 1 & 2 & -1 \\ 5 & 3 & 4 \\ -2 & 0 & 1 \end{vmatrix}
&= 1\begin{vmatrix} 3 & 4 \\ 0 & 1 \end{vmatrix} - 5\begin{vmatrix} 2 & -1 \\ 0 & 1 \end{vmatrix} + -2\begin{vmatrix} 2 & -1 \\ 3 & 4 \end{vmatrix} \\
&= \Big(3-0\Big) - 5\Big(2-0\Big) - 2\Big(8+3\Big) \\
&= 3 - 10 - 22 \\
&= -29
\end{aligned}
$$

The general, $n \times n$ case

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

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

(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

Bases and coefficients

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 are *linearly independent*:

$$a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n = \mathbf{0} \quad \Longrightarrow \quad \text{all} \;\; a_i = 0$$

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

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

Bases, equivalently

Equivalently: 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:

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

It's useful to think of column vectors just as notation for this sum:

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

Previously, we haven't bothered to write $\mathcal{B}$, but it is important!

Example: two bases for $\mathbb{R}^2$

Let $V = \mathbb{R}^2$, and let $\mathcal{S} = \{(1,0),(0,1)\}$ be the *standard basis*.

Vectors expressed in the standard basis give exactly what you expect:

$$\begin{pmatrix} a \\ b \end{pmatrix}_{\mathcal{S}} = a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

But expressing a vector in another basis can give something totally different! For example, let $\mathcal{B} = \{(100,0),(100,1)\}$:

$$\begin{pmatrix} a \\ b \end{pmatrix}_{\mathcal{B}} = a \cdot \begin{pmatrix} 100 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 100 \\ 1 \end{pmatrix} = \begin{pmatrix} 100 \cdot (a+b) \\ b \end{pmatrix}$$

Same vector, different outfits

Hence 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:

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

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

Why???

- Many find the idea of *multiple bases* confusing the first time around.
- $\mathcal{S} = \{(1,0),(0,1)\}$ is a perfectly good basis for $\mathbb{R}^2$. Why bother with others?
  1. Some vector spaces don't have one "obvious" choice of basis. Example: subspaces $S \subseteq \mathbb{R}^n$.
  2. Sometimes it is way more efficient to write a vector with respect to a different basis, e.g.:

$$\begin{pmatrix} 93718234 \\ -438203 \\ 110224 \\ -5423204980 \\ \vdots \end{pmatrix}_{\mathcal{S}} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}_{\mathcal{B}}$$

  3. The choice of basis for *vectors* affects how we write *matrices* as well. Often this can be done cleverly. Example: JPEGs, Google

Transforming bases, part I

- How can we transform a vector form the standard basis to a new basis, e.g. $\mathcal{B} = \{(100, 0), (100, 1)\}$?
- In order to express $(a, b) \in \mathbb{R}^2$ in $\mathcal{B}$ we need to find $x, y \in \mathbb{R}$ such that:

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

- Solving the equations gives: $y = b$ and $x = \frac{a-100b}{100}$

Example

The vector $\boldsymbol{v} = (100, 10) \in \mathbb{R}^2$ is represented w.r.t. the basis $\mathcal{B}$ as:

$$\begin{pmatrix} -9 \\ 10 \end{pmatrix}_{\mathcal{B}} = -9 \cdot \begin{pmatrix} 100 \\ 0 \end{pmatrix} + 10 \cdot \begin{pmatrix} 100 \\ 1 \end{pmatrix} = \begin{pmatrix} 100 \\ 10 \end{pmatrix}_{\mathcal{S}}$$

(use $a = 100, b = 10$ in the formulas for x, y given above.)

Transforming bases, part II

- **Easier:** 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 III

- How do we go back? Need $\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}}$ which does this:

$$\begin{pmatrix} a \\ b \end{pmatrix}_{\mathcal{S}} \rightsquigarrow \begin{pmatrix} \frac{a-100b}{100} \\ b \end{pmatrix}_{\mathcal{B}}$$

- 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:

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

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

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:

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

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

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

- Then, by linearity:

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

On the other hand...

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

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

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

- Then, by linearity:

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

A new matrix

- From this:

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

- It follows that we should instead us *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 totally different way!

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

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

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

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

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 $\boldsymbol{T}_{\mathcal{B}\Rightarrow C}$ and $\boldsymbol{T}_{C\Rightarrow\mathcal{B}}$ 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

Example basis transformation, part II

- The basis transformation matrix $\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}}$ in the other direction is obtained as matrix inverse:

$$\boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}} = \left(\boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}}\right)^{-1} = \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix}^{-1} = \frac{1}{-2-0}\begin{pmatrix} 2 & 0 \\ -1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix}$$

- Hence:

$$\begin{aligned} \boldsymbol{A}' &= \boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}} \cdot \boldsymbol{A} \cdot \boldsymbol{T}_{\mathcal{B}\Rightarrow\mathcal{S}} \\ &= \frac{1}{2}\begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} -2 & 2 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 4 & 4 \\ -1 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 2 \\ -\frac{1}{2} & 2 \end{pmatrix} \end{aligned}$$

Example basis transformation, part III

- Consider a vector $\boldsymbol{v} \in \mathbb{R}^2$ which can be represented in bases $\mathcal{S}$ and $\mathcal{B}$ respectively as:

$$\begin{pmatrix} 5 \\ 4 \end{pmatrix}_{\mathcal{S}} \quad \text{and} \quad \begin{pmatrix} -5 \\ 4\frac{1}{2} \end{pmatrix}_{\mathcal{B}}$$

- That is, we have:

$$\boldsymbol{v}' := \boldsymbol{T}_{\mathcal{S}\Rightarrow\mathcal{B}} \cdot \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} -5 \\ 4\frac{1}{2} \end{pmatrix}$$

- Then, if we apply $\boldsymbol{A} = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$ to $\boldsymbol{v}$, written in the basis $\mathcal{S}$, we get:

$$\boldsymbol{A} \cdot \boldsymbol{v} = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 22 \end{pmatrix}$$

Example basis transformation, part IV

- On the other hand, if we apply $\boldsymbol{A}' = \begin{pmatrix} 2 & 2 \\ -\frac{1}{2} & 2 \end{pmatrix}$ to $\boldsymbol{v}'$ we get:

$$\boldsymbol{A}' \cdot \boldsymbol{v}' \;=\; \begin{pmatrix} 2 & 2 \\ -\frac{1}{2} & 2 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 4\frac{1}{2} \end{pmatrix} \;=\; \begin{pmatrix} -1 \\ 11\frac{1}{2} \end{pmatrix}$$

  ...which we interpret as a vector written in $\mathcal{B}$.

- Comparing the two results:

$$\begin{pmatrix} -1 \\ 11\frac{1}{2} \end{pmatrix}_{\mathcal{B}} = -1 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} + 11\tfrac{1}{2} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 22 \end{pmatrix} = \begin{pmatrix} 1 \\ 22 \end{pmatrix}_{\mathcal{S}}$$

...we get the same outcome!

In fact: this is always the case. It can be shown using the definitions of $\boldsymbol{A}'$, $\boldsymbol{v}'$ and properties of inverses (i.e. no *matrixrekenen* necessary!).