Lecture 15:
Review
Change of basis

Notation cheat sheet

notation	meaning
F	Field with $0 \neq 1$
U, V, W	Vector spaces over a field F
$M_{\ell,k}(F)$	Matrices with ℓ rows and k columns, with entries in F *Note:* $M_k(F) = M_{k,k}(F)$
$\mathcal{B} = \langle \mathbf{b}_1, \ldots, \mathbf{b}_\ell \rangle$	Ordered basis
$\mathcal{E} = \langle \mathbf{e}_1, \ldots, \mathbf{e}_\ell \rangle$	Standard (ordered) basis of F^ℓ
$\mathcal{E} = \langle E_{1,1}, E_{1,2}, \ldots, E_{\ell,k} \rangle$	Standard (ordered) basis of $M_{\ell,k}(F)$
$\dim_F(V) = \dim(V)$	Dimension of V as a vector space over F—the size of any basis of V.
$\mathrm{Rep}_{\mathcal{B}}(\mathbf{v})$	The vector representation of $\mathbf{v} \in V$ with respect to an ordered basis $\mathcal{B}$ (it is an element of $F^{\dim(V)}$).
$\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(f)$	The matrix representation of $f : U \to V$ w.r.t ordered bases $\mathcal{A}$ and $\mathcal{B}$ of U and V respectively (it is an element of $M_{\dim(V),\dim(U)}(F)$).
$\mathbf{u} \cdot \mathbf{v}$	Dot product: $(u_1, \ldots, u_\ell) \cdot (v_1, \ldots, v_\ell) = u_1 v_1 + \cdots + u_\ell v_\ell$ $\mathbf{u} \cdot \mathbf{v} = \lvert\mathbf{u}\rvert\lvert\mathbf{v}\rvert \cos(\theta)$

Where *were* we??

Linear functions between f.d. vectors spaces are "the same" as matrices.

The fine print: Let U, V be vector spaces over a field F with $\dim(U) = k$ and $\dim(V) = \ell$. Let $\mathcal{A} = \langle \mathbf{a}_1, \ldots, \mathbf{a}_k \rangle$ be a basis of U and let $\mathcal{B} = \langle \mathbf{b}_1, \ldots, \mathbf{b}_\ell \rangle$ be a basis of V. Let $f : U \to V$ be a linear function. Define

$$\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(f) = \begin{pmatrix} | & | & & | \\ \mathbf{col}_1 & \mathbf{col}_2 & \cdots & \mathbf{col}_k \\ | & | & & | \end{pmatrix} \in M_{\ell,k}(F) \quad \text{where} \quad \mathbf{col}_i = \mathrm{Rep}_{\mathcal{B}}\, \text{f(a\_i)} \in F^\ell.$$

Then $\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}} : \mathrm{Hom}(U, V) \to M_{\ell,k}(F)$ is an isomorphism;

and for all $\mathbf{u} \in U$, we have $\boxed{\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(f)\mathrm{Rep}_{\mathcal{A}}(\mathbf{u}) = \mathrm{Rep}_{\mathcal{B}}(f(\mathbf{u}))}.$

⚠ **Caution!** Even if $\mathbf{u}$ is already a vector in F^k, unless $\mathcal{A} = \mathcal{E}$, you need to expand/represent $\mathbf{u}$ in the basis $\mathcal{A}$ *before* multiplying it by the matrix.

Matrix multiplication corresponds to function composition.

The fine print: We defined the product of matrices $X \in M_{m,\ell}(F)$ and $Y \in M_{\ell,k}$ as the matrix $XY \in M_{m,k}(F)$ whose (i, j) entry is $\mathbf{row}_i(X) \cdot \mathbf{col}_j(Y)^T$. Then for any functions $f : U \to V$ and $g : V \to W$, we have

$$\mathrm{Rep}_{\mathcal{B}}^{\mathcal{C}}(g)\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(f) = \mathrm{Rep}_{\mathcal{A}}^{\mathcal{C}}(g \circ f),$$

where $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are ordered bases of U, V, and W respectively.

[Where *were* we?? continued...]

"Standard" matrices are great building blocks

The fine print: In $M_{\ell,k}(F)$, we defined $E_{i,j}$ as the $\ell \times k$ matrix that has a 1 in row i and col j and 0's elsewhere. This means it encodes the function that sends $\mathbf{e}_j \mapsto \mathbf{e}_i$ and $\mathbf{e}_r \mapsto 0$ for all $r \neq j$: $\boxed{E_{i,j}\mathbf{e}_r = \delta_{j,r}\mathbf{e}_i}$. The set $\mathcal{E} = \{E_{i,j} \mid 1 \leqslant i \leqslant \ell, 1 \leqslant j \leqslant k\}$ is the standard basis of $M_{\ell,k}(F)$. Moreover, for any $X \in M_{m,\ell}(F)$, we saw that

$E_{i,j}X$ has $\mathbf{row}_j(X)$ in row i and 0's elsewhere, and $\quad (*)$

$XE_{i,j}$ has $\mathbf{col}_i(X)$ in col j and 0's elsewhere. $\quad (**)$

HW: Use the identities $E_{i,j}\mathbf{e}_r = \delta_{j,r}\mathbf{e}_i$ and $X = \sum_{i,j} X_{i,j}E_{i,j}$ to prove $\boxed{E_{i,j}E_{r,s} = \delta_{j,r}E_{i,s}}$, followed by $(*)$ and $(**)$.

Identity matrix. The $\ell \times \ell$ matrix corresponding to the identity map has 1s on the main diagonal and 0s elsewhere.

$$I_\ell = \mathrm{Rep}_{\mathcal{B}}^{\mathcal{B}}(\mathrm{id}_V) = \sum_{i=1}^{\dim(V)} E_{i,i} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

[Where *were* we?? continued...]

We can use row operations to invert matrices

The fine print: Row operations can be done by left multiplication by **elementary reduction matrices**.

$$S_i(\lambda) = \lambda E_{i,i} + \sum_{\substack{r=1,\dots,\ell \\ r\neq i}} E_{r,r} \qquad \text{"scale"}$$

$$P_{i,j} = E_{i,j} + E_{j,i} + \sum_{\substack{r=1,\dots,\ell \\ r\neq i,j}} E_{r,r} \qquad \text{"permute"}$$

$$C_{i,j}(\lambda) = \lambda E_{j,i} + I_\ell \qquad \text{"combine"}$$

If X row reduces to the identity matrix by operations

$$X \xmapsto{R_1} X_2 \xmapsto{R_2} X_3 \xmapsto{R_3} X_4 \cdots \xmapsto{R_n} I_\ell$$

where R_i are elementary red. matrices, then

$$R_n \cdots R_3 R_2 R_1 X = I_\ell, \quad \text{so that} \quad R_n \cdots R_3 R_2 R_1 = X^{-1}.$$

Note: This says that if I take the *exact same sequence* of operations that I did to move from X to I_ℓ, but instead I do them to I_ℓ, I will end with X^{-1}:

$$R_n \cdots R_3 R_2 R_1 I_\ell = X^{-1}.$$

Example. To compute the inverse of $X = \begin{pmatrix} 1 & -3 \\ -5 & 12 \end{pmatrix}$, we start by row reducing:

$$\begin{pmatrix} 1 & -3 \\ -5 & 12 \end{pmatrix} \xmapsto[C_{1,2}(5)]{\mathbf{row}_2 \mapsto \mathbf{row}_2 + 5\mathbf{row}_1} \begin{pmatrix} 1 & -3 \\ 0 & -3 \end{pmatrix}$$

$$\xmapsto[C_{2,1}(-1)]{\mathbf{row}_1 \mapsto \mathbf{row}_1 - \mathbf{row}_2} \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \xmapsto[S_2(-1/3)]{\mathbf{row}_2 \mapsto (-1/3)\mathbf{row}_2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$X^{-1} = S_2(-1/3) C_{2,1}(-1) C_{1,2}(5) = \begin{pmatrix} 1 & 0 \\ 0 & -1/3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} -4 & -1 \\ -5/3 & -1/3 \end{pmatrix}.$$

Shortcut for doing calculations by hand.

Augment X by I_ℓ and *then* row reduce:

$$\left(\begin{array}{cc|cc} 1 & -3 & 1 & 0 \\ -5 & 12 & 0 & 1 \end{array}\right) \xmapsto{\mathbf{row}_2 \mapsto \mathbf{row}_2 + 5\mathbf{row}_1} \left(\begin{array}{cc|cc} 1 & -3 & 1 & 0 \\ 0 & -3 & 5 & 1 \end{array}\right)$$

$$\xmapsto{\mathbf{row}_1 \mapsto \mathbf{row}_1 - \mathbf{row}_2} \left(\begin{array}{cc|cc} 1 & 0 & -4 & -1 \\ 0 & -3 & 5 & 1 \end{array}\right)$$

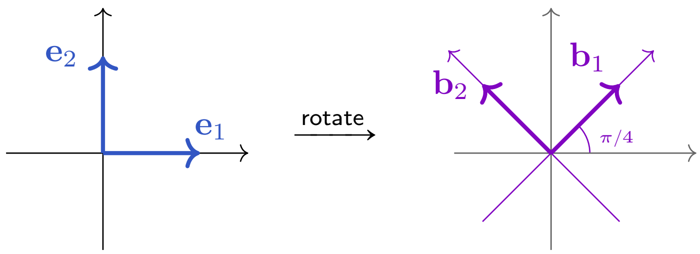
$$\xmapsto{\mathbf{row}_2 \mapsto (-1/3)\mathbf{row}_2} \left(\begin{array}{cc|cc} 1 & 0 & -4 & -1 \\ 0 & 1 & -5/3 & -1/3 \end{array}\right).$$

Change of basis

Suppose we computed a matrix for a linear function $f : U \to V$ with respect to ordered bases $\mathcal{A}$ (of U) and $\mathcal{B}$ (of V), but what we *want* is the linear function in terms of a *different* set of ordered bases $\mathcal{A}'$ (of U) and $\mathcal{B}'$ (of V)?

Examples.

1. If $f : F^2 \to F^2$ sends $\begin{pmatrix}1\\3\end{pmatrix} \mapsto \begin{pmatrix}5\\0\end{pmatrix}$ and $\begin{pmatrix}0\\-2\end{pmatrix} \mapsto \begin{pmatrix}1\\1\end{pmatrix}$. What is $\mathrm{Rep}_{\mathcal{E}}^{\mathcal{E}}(f)$?
2. Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $\mathrm{Rep}_{\mathcal{E}}^{\mathcal{E}}(f) = \begin{pmatrix}1 & 2\\2 & 1\end{pmatrix}$. What is the matrix for f with respect to a frame of reference rotated by $\pi/4$?



3. How does the evaluation map compare on polynomials "Taylor expanded" around $a = 0$ versus around $a = 1$?

Big idea: For any functions $f : U \to V$ and $g : V \to W$, we have

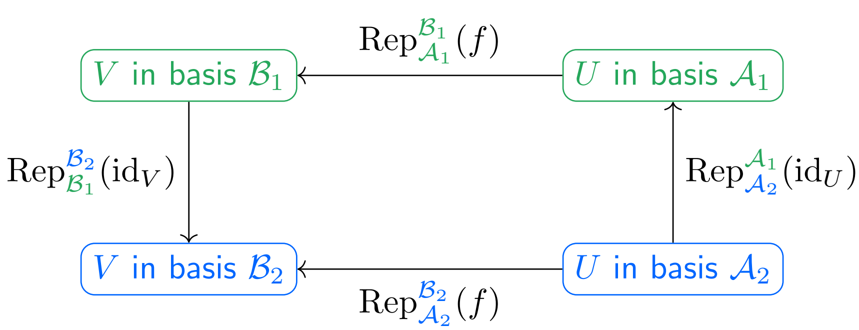
$$\boxed{\mathrm{Rep}_{\mathcal{B}}^{\mathcal{C}}(g)\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(f) = \mathrm{Rep}_{\mathcal{A}}^{\mathcal{C}}(g \circ f)}$$

where $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are ordered bases of U, V, and W respectively.

So for ordered bases $\mathcal{A}_1$ and $\mathcal{A}_2$ of U and ordered bases $\mathcal{B}_1$ and $\mathcal{B}_2$ of V, we have

$$\mathrm{Rep}_{\mathcal{B}_1}^{\mathcal{B}_2}(\mathrm{id}_V)\mathrm{Rep}_{\mathcal{A}_1}^{\mathcal{B}_1}(f)\mathrm{Rep}_{\mathcal{A}_2}^{\mathcal{A}_1}(\mathrm{id}_U) = \mathrm{Rep}_{\mathcal{A}_2}^{\mathcal{B}_2}(f),$$

since $\mathrm{id}_V \circ f \circ \mathrm{id}_U = f$.



Example: If $f : F^2 \to F^2$ sends $\begin{pmatrix}1\\3\end{pmatrix} \mapsto \begin{pmatrix}5\\0\end{pmatrix}$ and $\begin{pmatrix}0\\-2\end{pmatrix} \mapsto \begin{pmatrix}1\\1\end{pmatrix}$.

What is $\mathrm{Rep}_{\mathcal{E}}^{\mathcal{E}}(f)$?

Ans. Let $\mathcal{B} = \left\langle \mathbf{b}_1 = \begin{pmatrix}1\\3\end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix}0\\-2\end{pmatrix} \right\rangle$.

Then

$$\mathbf{b}_1 = \begin{pmatrix}1\\3\end{pmatrix} \stackrel{f}{\longmapsto} \begin{pmatrix}5\\0\end{pmatrix} = 5\mathbf{e}_1 + 0\mathbf{e}_2 \quad \text{and} \quad \mathbf{b}_2 = \begin{pmatrix}0\\-2\end{pmatrix} \stackrel{f}{\longmapsto} \begin{pmatrix}1\\1\end{pmatrix} = 1\mathbf{e}_1 + 1\mathbf{e}_2$$

means

$$\mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(f) = \begin{pmatrix}5 & 1\\0 & 1\end{pmatrix}. \qquad \text{But we want} \quad \mathrm{Rep}_{\mathcal{E}}^{\mathcal{E}}(f) = \mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(f)\mathrm{Rep}_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}).$$

Computing $\mathrm{Rep}_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id})$**...**

Old perspective: Our job would be to solve

$$\mathbf{e}_1 = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 \quad \text{and} \quad \mathbf{e}_2 = d_1\mathbf{b}_1 + d_2\mathbf{b}_2 \quad \text{for } c_1, c_2, d_1, d_2 \in F.$$

Row reduce

$$\left(\begin{array}{cc|c}1 & 0 & 1\\3 & -2 & 0\end{array}\right) \quad \text{and} \quad \left(\begin{array}{cc|c}1 & 0 & 0\\3 & -2 & 1\end{array}\right)$$

New perspective: Note that

$$\mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id})\mathrm{Rep}_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}) = \mathrm{Rep}_{\mathcal{E}}^{\mathcal{E}}(\mathrm{id} \circ \mathrm{id}) = I_2.$$

So

$$\boxed{\mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}) = \left(\mathrm{Rep}_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id})\right)^{-1}.}$$

But

$$\begin{array}{l}\mathbf{b}_1 \stackrel{\mathrm{id}}{\longmapsto} \mathbf{b}_1 = \mathbf{e}_1 + 3\mathbf{e}_2 \quad \text{and}\\ \mathbf{b}_2 \stackrel{\mathrm{id}}{\longmapsto} \mathbf{b}_2 = 0\mathbf{e}_1 + (-2)\mathbf{e}_2\end{array} \quad \text{means} \quad \mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}) = \begin{pmatrix}1 & 0\\3 & -2\end{pmatrix}.$$

Row reduce:

$$\left(\begin{array}{cc|cc}1 & 0 & 1 & 0\\3 & -2 & 0 & 1\end{array}\right) \xrightarrow{\mathbf{row}_2 \mapsto \mathbf{row}_2 - 3\mathbf{row}_1} \left(\begin{array}{cc|cc}1 & 0 & 1 & 0\\0 & -2 & -3 & 1\end{array}\right)$$

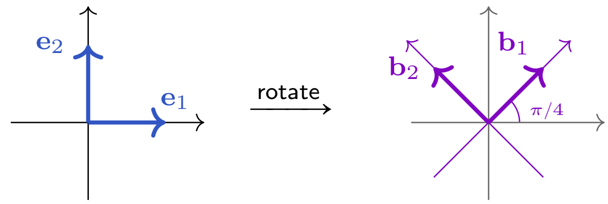
$$\xrightarrow{\mathbf{row}_2 \mapsto (-1/2)\mathbf{row}_2} \left(\begin{array}{cc|cc}1 & 0 & 1 & 0\\0 & 1 & 3/2 & -1/2\end{array}\right)$$

Therefore $\mathrm{Rep}_{\mathrm{E}}^{\mathrm{B}}(\mathrm{id}) = \begin{pmatrix}1 & 0\\3/2 & -1/2\end{pmatrix}$, and hence

$$\mathrm{Rep}_{\mathcal{E}}^{\mathcal{E}}(f) = \mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(f)\mathrm{Rep}_{\mathcal{E}}^{\mathcal{B}}(\mathrm{id}) = \begin{pmatrix}5 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\3/2 & -1/2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}13 & -1\\3 & -1\end{pmatrix}.$$

You try:

1. Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $\text{Rep}^{\mathcal{E}}_{\mathcal{E}}(f) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. What is the matrix for f with respect to a frame of reference rotated by $\pi/4$?



   (a) Compute $\mathbf{b}_1$ and $\mathbf{b}_2$ in terms of $\mathbf{e}_1$ and $\mathbf{e}_2$.
   (b) Use (a) to write down $\text{Rep}^{\mathcal{E}}_{\mathcal{B}}(\text{id})$.
   (c) Compute $\text{Rep}^{\mathcal{B}}_{\mathcal{E}}(\text{id}) = (\text{Rep}^{\mathcal{E}}_{\mathcal{B}}(\text{id}))^{-1}$.
   (d) Use (b) and (c) to compute $\text{Rep}^{\mathcal{B}}_{\mathcal{B}}(f)$.

2. Let $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some $a, b, c, d \in F$.

   (a) Row reduce $\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right)$ to compute X^{-1} for a general 2×2 matrix. Keep track of when you might be accidentally dividing by 0.
   (b) What does it mean if you can't row reduce $(X|I)$ without dividing by 0? Illustrate with an example.
   (c) Compare your formula with our computations of inverses of 2×2 matrices thus far in this lecture.

Let $\mathcal{A}$ and $\mathcal{B}$ be ordered bases of a vector space V. We call $\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(\mathrm{id})$ the **change of basis matrix** from $\mathcal{A}$ to $\mathcal{B}$.

Because it's worth highlighting...
Lemma. We have

$$\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(\mathrm{id}) = \left(\mathrm{Rep}_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id})\right)^{-1}.$$

Consequences:

1. If P is a change of basis matrix, then P is invertible.
2. If P is an invertible matrix, it encodes an isomorphism. Thus the image of the basis $\mathcal{E}$, $\mathcal{B} = \langle P\mathbf{e}_i \mid \mathbf{e}_i \in \mathcal{E}\rangle$, is also a basis. So

$$P = \begin{pmatrix} | & | & & | \\ P\mathbf{e}_1 & P\mathbf{e}_2 & \cdots & P\mathbf{e}_\ell \\ | & | & & | \end{pmatrix} = \mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}).$$

Theorem. The set of invertible $\ell \times \ell$ matrices is the same as the set of change of basis matrices.

Namely, two square matrices $X, Y \in M_\ell(F)$ **represent the same function** $f : V \to V$ (where $\dim(V) = \ell$), but with respect to different bases, if and only if there is some invertible $P \in M_\ell(F)$ for which $Y = PXP^{-1}$; in this case we say X and Y are **conjugate** or **similar**. We call $X \mapsto PXP^{-1}$ **conjugating** X by P.

"Similarity" defines an equivalence relation on the set $M_\ell(F)$:

$$X \sim Y \qquad \text{whenever} \qquad Y = PXP^{-1}$$

for some invertible $P \in M_\ell(F)$.

- Reflexive: If $X \in M_\ell(F)$, then...
- Symmetric: If $X, Y \in M_\ell(F)$ satisfy $X \sim Y$, then...
- Transitive: If $X, Y, Z \in M_\ell(F)$ satisfy $X \sim Y$ and $Y \sim Z$, then...

We'll be studying statistics about matrices that are invariant under change of basis, meaning that they're constant on similar matrices—these statistics are important because they pertain to the underlying functions independent of your choice of basis.