Lecture: Change of Basis

Date: October 8th, 2025 *Author: Surbhi Goel*

1 Recap and Preview

Last time we covered linear independence, basis, rank, and linear mappings. We saw that any linear mapping between finite dimensional vector spaces can be represented as a matrix once we choose bases. We also answered our motivating question: the system $X\theta = Y$ has a solution if and only if $\text{rk}(X) = \text{rk}(X|Y)$.

Today we explore a powerful idea: the *choice of basis matters*. The same linear transformation can look very different (and be much easier to work with) depending on which basis we use.

2 Change of Basis

Recall from last time: any vector $x \in V$ can be represented by its coordinate vector $\hat{x} \in \mathbb{R}^n$ once we choose a basis, and linear mappings can be represented as transformation matrices. But bases are not unique. The *same* linear transformation can have very different matrix representations depending on which basis we use.

Motivating example. Consider the linear transformation $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ with transformation matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

in the standard basis. If instead we use the basis $B = \{(1, 1), (1, -1)\}$, then Φ has transformation matrix

$$\tilde{A} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

which is diagonal (much nicer!). How do we know this is true? We'll develop the theory to answer this, then verify the claim.

2.1 The Basis Change Theorem

- For a linear mapping $\Phi : V \to W$, consider two bases $B = (b_1, \ldots, b_n)$ and $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ on V and two bases $C = (c_1, \ldots, c_m)$ and $\tilde{C} = (\tilde{c}_1, \ldots, \tilde{c}_m)$ on W.
- Let $A_\Phi \in \mathbb{R}^{m \times n}$ be the transformation matrix of Φ with respect to B, C and let $\tilde{A}_\Phi$ be the transformation matrix of Φ with respect to $\tilde{B}, \tilde{C}$.
- How are A and $\tilde{A}$ related?

- Theorem: (Basis Change) The transformation matrix $\tilde{A}_\Phi$ is given by

$$\tilde{A}_\Phi = T^{-1} A_\Phi S$$

where $S \in \mathbb{R}^{n \times n}$ is the transformation matrix representing $\mathrm{id}_V : V \to V$ with respect to bases $\tilde{B}$ (input) and B (output), and $T \in \mathbb{R}^{m \times m}$ is the transformation matrix representing $\mathrm{id}_W : W \to W$ with respect to bases $\tilde{C}$ (input) and C (output).

- Proof: First, by definition of S we can write the $\tilde{b}_j$ as a sum of basis vectors b_i:

$$\tilde{b}_j = \sum_i s_{ij} b_i$$

Similarly, we can write $\tilde{c}_k$ as a combination of basis vectors of C:

$$\tilde{c}_k = \sum_l t_{lk} c_l$$

Then, S maps $\tilde{B}$ onto B and T maps $\tilde{C}$ onto C (the columns are the coordinate representation of $\tilde{b}_j$ and $\tilde{c}_k$ with respect to B and C). Now, re-express $\Phi(\tilde{b}_j)$ in two ways using these two bases. First using C:

$$\Phi(\tilde{b}_j) = \sum_{k=1}^{m} \tilde{a}_{kj} c_k = \sum_{k=1}^{m} \tilde{a}_{kj} \sum_{l=1}^{m} t_{lk} c_l = \sum_{l=1}^{m} c_l \sum_{k=1}^{m} t_{lk} \tilde{a}_{kj}$$

Then using B:

$$\Phi(\tilde{b}_j) = \Phi\left(\sum_{i=1}^{n} s_{ij} b_i\right) = \sum_{i=1}^{n} s_{ij} \Phi\left(b_i\right) = \sum_{i=1}^{n} s_{ij} \sum_{l=1}^{m} a_{li} c_l = \sum_{l=1}^{m} c_l \sum_{i=1}^{n} a_{li} s_{ij}$$

Therefore for all $j = 1, \ldots, n$ and all $l = 1, \ldots, m$ it follows that

$$\sum_{k=1}^{m} t_{lk} \tilde{a}_{kj} = \sum_{i=1}^{n} a_{li} s_{ij}$$

In matrix form, this is equivalent to

$$T\tilde{A} = AS$$

and therefore $\tilde{A} = T^{-1} A S$

- In particular, rank is basis-invariant: $\mathrm{rk}(T^{-1}AS) = \mathrm{rk}(A)$.

- Aside: Why are S and T regular (invertible)? They are the matrix representation of the identity operator, which is an invertible operator.

- Two matrices $A, \tilde{A} \in \mathbb{R}^{m \times n}$ are *equivalent* if there exists regular matrices $S \in \mathbb{R}^{n \times n}, T \in \mathbb{R}^{m \times m}$ such that $\tilde{A} = T^{-1} A S$

- Two matrices $A, \tilde{A} \in \mathbb{R}^{n \times n}$ are *similar* if there exists a regular matrix $S \in \mathbb{R}^{n \times n}$ where $\tilde{A} = S^{-1} A S$

- Informally, this basis change can be seen as the following:

- A maps $V \to W$ bases B to C
- $\tilde{A}$ maps $V \to W$ from bases $\tilde{B}$ to $\tilde{C}$
- S is the identity mapping from basis $\tilde{B}$ to B
- T is the identity mapping from basis $\tilde{C}$ to C
- Then, $\tilde{B} \to \tilde{C}$ can be rewritten as

$$\tilde{B} \to B \to C \to \tilde{C}$$

which reflects S, then A, then T^{-1}. Hence, $\tilde{A}x = T^{-1}(A(Sx))$

- **Computing the motivating example.** To get the diagonal matrix from the start: note that $S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ (just horizontally stack the new basis vectors) and that

$$T^{-1} = S^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then,

$$\tilde{A} = T^{-1}AS = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

This confirms our claim: the right basis makes the transformation diagonal!

3 Images and Kernels

Now that we understand how to change bases for linear mappings, we cover the fundamental concepts of images and kernels (which you've seen briefly as the column space and null space).

- For $\Phi : V \to W$ the *kernel* or null space is $\ker(\Phi) = \Phi^{-1}(0) = \{v \in V : \Phi(v) = 0\}$
- This is the set of vectors in V that map to 0
- The *image* or range is $\text{Im}(\Phi) = \Phi(V) = \{w \in W \mid \exists v \in V : \Phi(v) = w\}$
- This is the set of vectors in W that can be reached by Φ
-
  - It is always true that $\Phi(0) = 0$ and therefore $0 \in \ker(\Phi)$, so the null space is never empty.
  - It is also always true that $\ker(\Phi) \subseteq V$ and $\text{Im}(\Phi) \subseteq W$.
  - Φ is injective if and only if $\ker(\Phi) = \{0\}$
- Consider a linear mapping $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ with transformation matrix $A \in \mathbb{R}^{m \times n}$, so $\Phi(x) = Ax$
- Let $A = [a_1, \ldots, a_n]$ be the columns of A. Then the image is the span of the columns (column space):

$$\text{Im}(\Phi) = \{Ax : x \in \mathbb{R}^n\} = \left\{ \sum_i x_i a_i : x_i \in \mathbb{R} \right\} = \text{span}(a_1, \ldots, a_n) \subseteq \mathbb{R}^m$$

- Then it follows that the rank of A is the dimension of the image, i.e. $\text{rk}(A) = \dim(\text{Im}(\Phi))$
- Rank Nullity Theorem: For vector spaces V, W and linear mapping $\Phi : V \to W$ it holds that

$$\dim(\ker(\Phi)) + \dim(\text{Im}(\Phi)) = \dim(V)$$

  also known as the fundamental theorem of linear mappings.
- Some immediate consequences:
  - If $\dim(\text{Im}(\Phi)) < \dim(V)$ then $\ker(\Phi)$ is non-trivial
  - If A_Φ is the transformation matrix of Φ and $\dim(\text{Im}(\Phi)) < \dim(V)$ then $Ax = 0$ has infinitely many solutions
  - If $\dim(V) = \dim(W)$ then Φ is injective if and only if it is surjective

4 Affine Subspaces

The last part we will consider here is affine subspaces. These are subspaces that have a linear structure.

- Let V be a vector space $x_0 \in V$ and $U \subseteq V$ be a subspace. Then

$$L = x_0 + U = \{x_0 + u : u \in U\}$$

  is an affine subspace.
- Examples of affine subspaces: points, lines, planes...
- If $(b_1, \ldots, b_k)$ is an ordered basis of U then every element $x \in L$ is uniquely described as $x = x_0 + \sum_i \lambda_i b_i$
- In the same way that we can define linear mappings between vector spaces, we can also define affine mappings between affine subspaces.
- For two vector spaces V, W, linear mapping $\Phi : V \to W$ and $a \in W$, the mapping

$$\phi : V \to W$$

$$x \to a + \Phi(x)$$

  is an affine mapping from V to W, where a is the translation vector.
- Every affine mapping is the composition of a linear mapping Φ and a translation τ such that $\phi = \tau \circ \Phi$
- Composition $\phi' \circ \phi$ of affine operators is affine
- Affine operators preserve dimension and parallelism and other geometric structures