

Math 221 Lec 16

4.3: Linear transformations & change of basis

Asa Royal (ajr74)

11/2/23, 11/9/23

Definition 1 (linear transformation). A function $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is called a **linear transformation** if it satisfies

1. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
2. $T(c\mathbf{x}) = cT(\mathbf{x}) \forall \mathbf{x} \in \mathbb{R}^n$ and scalars c

Remark. If $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is a linear transformation, then we can find a matrix A , the so-called standard matrix of A so that $T = \mu_A$. The j th column of A is given by $T(\mathbf{e}_j)$, where \mathbf{e}_j is the j th standard basis vector.

Proposition 2. Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation and let $\mathcal{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis for \mathbb{R}^n . Let A be the matrix whose column vectors are the vectors $T(\mathbf{e}_1), \dots, T(\mathbf{e}_n) \in \mathbb{R}^m$ (that is, the coordinate vectors of $T(\mathbf{e}_j)$ with respect to the standard basis of \mathbb{R}^m):

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e})_1 & T(\mathbf{e})_2 & \dots & T(\mathbf{e})_n \\ | & | & & | \end{bmatrix}$$

Then $T = \mu_A$ and we call A the standard matrix for T , denoted $[T]_{\text{stand}}$.

Proof.

□

Definition 3 (matrix with respect to basis). Let $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ be a linear transformation and let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an ordered basis for \mathbb{R}^n . For each $j = 1, \dots, n$, let $a_{1j}, a_{2j}, \dots, a_{nj}$ denote the coordinates of $T(\mathbf{v}_j)$ with respect to the basis \mathcal{B} . We denote this matrix by $[T]_{\mathcal{B}}$.

Remark. The coefficients of $T(\mathbf{v}_j)$ will be entered as the j th column of the matrix $[T]_{\mathcal{B}}$. I.e. given a vector $\mathbf{x} \in \mathbb{R}^n$, we let $C_{\mathcal{B}(\mathbf{x})}$ denote the column vector whose entries are the coordinates of \mathbf{x} w.r.t. the basis of \mathcal{B} . That is, if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for \mathcal{B} ,

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$$

then

$$C_{\mathcal{B}(\mathbf{x})} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

So

$$[T]_{\mathcal{B}} = \begin{bmatrix} | & | & & | \\ C_{\mathcal{B}(T(\mathbf{v}_1))} & C_{\mathcal{B}(T(\mathbf{v}_2))} & \dots & C_{\mathcal{B}(T(\mathbf{v}_n))} \\ | & | & & | \end{bmatrix}$$

Definition 4 (coordinates w.r.t. basis). Let V be a subspace of \mathbb{R}^n and let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a basis for V . Then for $\mathbf{v} \in V$, $\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k$. c_1, \dots, c_k are the coordinates of \mathbf{v} with respect to basis \mathcal{B} .

Example. Say $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for \mathbb{R}^2 with $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Further, let's consider the vector $\mathbf{a} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$. Note that $\mathbf{a} = 3\mathbf{v}_1 + 2\mathbf{v}_2$. So $[\mathbf{a}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. That is, the coordinates for \mathbf{a} with respect to the basis \mathcal{B} are $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

Remark. When we talk offhand about Cartesian coordinates $(2, 1)$ for example, we are usually referring to those as coordinates with respect to the standard basis for \mathbb{R}^2 , $[\mathbf{e}_1 \ \mathbf{e}_2]$. $(2, 1) = 2\mathbf{e}_1 + 1\mathbf{e}_2$.

Definition 5 (change-of-basis matrix). The change of basis matrix for some basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is simply $[\mathbf{v}_1 \ \dots \ \mathbf{v}_k]$. It represents the linear transformation of basis vectors in one basis to basis vectors in another.

Example. Imagine we have $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ in some basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ and we want to get its coordinates in the standard basis. Then $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ is the change of basis matrix between \mathcal{B} and ε , so we can take $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and get coordinates w.r.t. the standard basis.

If we had $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ in standard basis coordinates and wanted to find its coordinates with respect to \mathcal{B} , We could take $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Remark. Can kinda think of the change of basis matrix as translating between coordinate systems (languages for representing vectors). The matrix with basis 2's basis vectors written in coordinates (language) of basis 1 shows how to translate basis 2's coordinates into basis 1 coordinates (language).

Remark. We can figure out coordinates for linear transformations in any basis through the following process. Imagine a matrix A that represents linear transformation T in some basis \mathcal{B}_{∞} , and we want to find the matrix $[T]_{\mathcal{B}_2}$, which tells us how T affects vectors with coordinates written w.r.t. basis \mathcal{B}_2 . First, we take some vector \mathbf{x} with coordinates in \mathcal{B}_1 and convert its coordinates to \mathcal{B}_1 . How? Change of basis matrix $P =$ basis vectors of \mathcal{B}_2 . So we take $P\mathbf{x}$. We then apply the linear transformation T , as represented by A in \mathcal{B}_{∞} , which gives us $AP\mathbf{x}$. Finally, we convert coordinates back to \mathcal{B}_2 . Then the coordinates for $T[\mathbf{x}]$ w.r.t. \mathcal{B}_2 are given by $P^{-1}AP$.

Remark. If $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ maps $\mathbf{x} \mapsto T(\mathbf{x})$, T of course maps $[\mathbf{x}]_{\mathcal{B}} \mapsto [T(\mathbf{x})]_{\mathcal{B}}$, since the vectors in the latter statement are the same ones as in the first, simply expressed w.r.t a different basis. But the matrix we use to represent T changes depending on the basis we're operating w.r.t. see below

Remark. Why would we transform bases? Some matrices are much easier to multiply in certain bases. For example, if we can transform a matrix into a basis where it's diagonal, we save a lot of multiplication computations.

Definition 6 (alt. def matrix w.r.t. basis). The linear transformation of basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a linear combination of of basis vectors $\mathbf{w}_1, \dots, \mathbf{w}_m$ (from the target basis) with coefficients (coordinates) given by the matrix $[T]_{\mathcal{V}, \mathcal{W}}$. The j th column of $[T]_{\mathcal{V}, \mathcal{W}}$ provides the coordinates for $T[\mathbf{v}_j]$ w.r.t. $\mathbf{w}_1, \dots, \mathbf{w}_m$

$$T \begin{bmatrix} \mathbf{v}_1 & \dots & \dots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1 & \dots & \mathbf{w}_m \end{bmatrix} [T]_{\mathcal{V}, \mathcal{W}} \quad (1)$$

Remark. Per the definition above, $T(\mathbf{v}_1) \in W$, which means $T\mathbf{v}_1$ is a linear combination of the combination of \mathbf{w}_i . Which

linear combination? The combination given by $T\mathbf{v}_1 = a_{11}\mathbf{w}_1 + \dots + a_{m1}\mathbf{w}_m = \begin{bmatrix} \mathbf{w}_1 & \dots & \mathbf{w}_m \end{bmatrix} \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}$

Remark. If we set $A = [T]_{\mathcal{V}, \mathcal{W}} \in \mathbb{R}^{m \times n}$, saying $A\mathbf{x} = \mathbf{b}$ means that the coefficients of \mathbf{x} of $\mathbf{v} \in V$ on $\mathbf{v}_1, \dots, \mathbf{v}_n$ get taken to the coefficients \mathbf{b} of $\mathbf{w} \in W$ on $\mathbf{w}_1, \dots, \mathbf{w}_m$ by $A = [T]_{\mathcal{V}, \mathcal{W}}$.

Lemma 7. Fix V with basis $\mathbf{v}_1, \dots, \mathbf{v}_n$. If $\mathbf{v}'_1, \dots, \mathbf{v}'_n$ is another basis of V then

$$\begin{bmatrix} \mathbf{v}'_1 & \dots & \mathbf{v}'_n \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix} P$$

for some invertible $P \in \mathbb{R}^{n \times n}$.

Remark. Why? Because each \mathbf{v}'_i can be expressed as a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Embed coefficients for the $\mathbf{v}_1, \dots, \mathbf{v}_n$ into the i th column of P . Note: We expect P to be invertible, because each \mathbf{v}_i can also be expressed as a linear combination of $\mathbf{v}'_1, \dots, \mathbf{v}'_n$.

Proof. The remark above serves as proof that P exists, but we can state more simply that $P = [\text{id}_V]_{\mathcal{V}', \mathcal{V}}$. I.e., P is the matrix for the identity linear transformation (that does absolutely nothing) with an input expressed w.r.t. \mathcal{V}' and output expressed w.r.t. \mathcal{V} . We can show P is invertible by showing that $\mathbf{x} \in N(P) \Rightarrow \mathbf{x} = \mathbf{0}$. If $\mathbf{x} \in N(P)$,

$$[\mathbf{v}'_1 \ \dots \ \mathbf{v}'_n] = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n] P \mathbf{x} = \mathbf{0}$$

Which means $x_1 \mathbf{v}'_1 + \dots + x_n \mathbf{v}'_n = \mathbf{0}$. But $\mathbf{v}'_1, \dots, \mathbf{v}'_n$ are linearly independent, so $\mathbf{x} = \mathbf{0}$ □

Definition 8 (change of basis matrix take 1000). P above is the change of basis matrix. It tells us how to express the basis vectors of V' in terms of the vectors of V . Again, see proof above—the change of basis matrix is the matrix for the identity linear transformation with an different input and output basis.

Definition 9 (change-of-basis formula). Fix

- a linear map $T : V \mapsto W$
- ordered bases
 $\mathcal{V}, \mathcal{V}'$ for V with $[\mathbf{v}'_1 \ \dots \ \mathbf{v}'_n] = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n] P$,
 $\mathcal{W}, \mathcal{W}'$ for W with $[\mathbf{w}'_1 \ \dots \ \mathbf{w}'_n] = [\mathbf{w}_1 \ \dots \ \mathbf{w}_n] Q$
- $A = [T]_{\mathcal{V}, \mathcal{W}}$, $A' = [T]_{\mathcal{V}', \mathcal{W}'}$

Then $A' = Q^{-1} A P$

Remark. A' is the matrix for the transform with respect to bases \mathcal{V}' and \mathcal{W}' , so we know that a formula for it starts by taking in a vector $\in V'$. It then changes bases from V' to V using P , applies the linear transform T , using coordinates from A so that the answer is w.r.t. basis W , then uses Q^{-1} to convert from bases W to W' . For a basis \mathcal{B} and $T : V \mapsto V$ (instead of $T : V \mapsto W$), we set $[T]_{\mathcal{B}} = [T]_{\mathcal{B}, \mathcal{B}}$. $[T]_{\mathcal{B}}$ takes the coefficients x of \mathbf{v} in V on $\mathbf{v}_1, \dots, \mathbf{v}_n$ to the coefficients of $T\mathbf{v}$ on $\mathbf{v}_1, \dots, \mathbf{v}_n$.