Lecture 6 - Rank nullity, change of basis

Monday, October 25, 2021 19:11

First lets formilize what we mean by a dimension of a linear subspace $V \subseteq \mathbb{R}^n$. For this, consider a

Theorem:

Number of vectors in a basis

All bases of a subspace V of $\mathbb{R}^n$ consist of the same number of vectors.

Proof: Exercize.*

Definition: Consider a subspace V of $\mathbb{R}^n$. The number of vectors in a basis of V is called the dimension of V, denoted by $\dim(V)$.

Examples: 1) Dimension of $\mathbb{R}^n$ is n, given, for example, by a basis $\vec{e}_1, \dots, \vec{e}_n$.

2) Dimension of a plane is 2, given by any two lin. indep. vectors in this plane.

3) Dimension of a line is 1.

4) $\dim(\operatorname{Im}(A)) = \operatorname{rank}(A)$, because we can construct a basis of $\operatorname{Im}(A)$ from column vectors in $\operatorname{rref}(A)$ with leading 1's.

Now, we come to one of the very important theorems in Linear Algebra.

Theorem (Rank-Nullity): Let $T: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a linear

Theorem (Rank-Nullity): Let $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a linear transformation, then

$$m = \dim(\mathbb{R}^m) = \dim(\ker T) + \dim(\operatorname{Im} T)$$

Proof: Choose a basis of $\operatorname{Im} T$, say $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_{\dim(\operatorname{Im} T)}$ and choose preimages $\vec{v}_1, \dots, \vec{v}_{\dim(\operatorname{Im} T)}$ so that

$$T(\vec{v}_i) = \vec{w}_i, \quad i = 1, \dots, \dim(\operatorname{Im} T)$$

Choose a basis of $\vec{u}_1, \dots, \vec{u}_{\dim(\ker T)}$ of $\ker T$. The result will follow once we show that

$$\vec{u}_1, \dots, \vec{u}_{\dim(\ker T)}, \vec{v}_1, \dots, \vec{v}_{\dim(\operatorname{Im} T)}$$

is a basis of $\mathbb{R}^m$.

Say $\vec{v} \in \mathbb{R}^m$. Since $T(\vec{v}) \in \operatorname{Im} T$, we have

$$T(\vec{v}) = \sum_{i=1}^{\dim(\operatorname{Im} T)} b_i \vec{w}_i, \text{ for some } b_i \in \mathbb{R}.$$

By linearity of T,

$$0 = T(\vec{v}) - T(\vec{v}) = \sum b_i \vec{w}_i - T(\vec{v}) = \sum b_i T(v_i) - T(\vec{v}) =$$
$$= T\left(\sum b_i v_i - v\right), \text{ so } \sum b_i v_i - v \in \ker T$$

Thus

$$\sum b_i v_i - v = \sum_{i=1}^{\dim(\ker T)} a_i u_i \quad \text{for some } a_i \in \mathbb{R}.$$

and

$$v = \sum b_i v_i - \sum a_i u_i, \text{ therefore}$$

$$v \in \operatorname{span}(\vec{u}_1, \dots, \vec{u}_{\dim(\ker T)}, \vec{v}_1, \dots, \vec{v}_{\dim(\operatorname{Im}(T))}).$$

Say now α_i, β_i are scalars such that

Say now, α_i, β_i are scalars such that

$$\sum \alpha_i u_i + \sum \beta_i v_i = 0 \qquad (*)$$

Then, applying T to both sides, we get that

$$\sum \beta_i w_i = 0$$

Since w_i are lin. independent, we get that

$$\beta_1 = \beta_2 = \cdots = \beta_{\dim(\operatorname{Im} T)} = 0$$

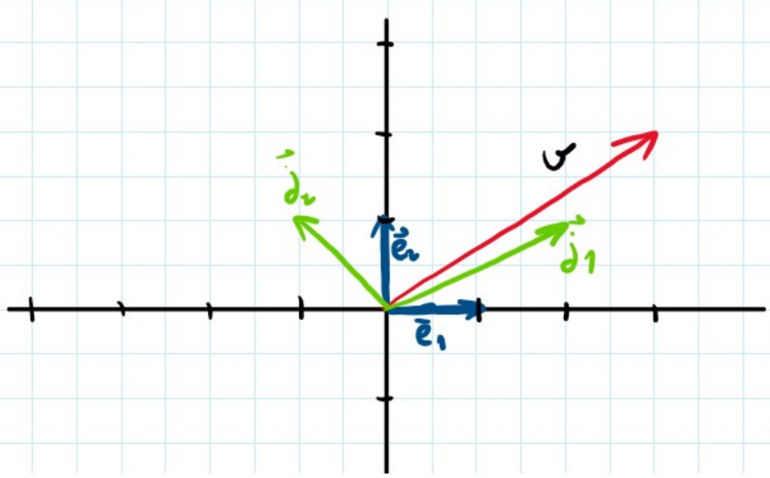
But, from $(*)$ we have now that $\sum \alpha_i u_i = 0$, and since u_i are also lin. indep. we get

$$\alpha_1 = \alpha_2 = \cdots = \alpha_{\dim(\ker T)} = 0.$$

Therefore

$u_1, \dots, u_{\dim(\ker T)}, v_1, \dots, v_{\dim(\operatorname{Im} T)}$ forms a basis of $\mathbb{R}^m$ $\square$.

Change of Basis.



Say we are given a vector $\vec{v} \in \mathbb{R}^2$, $\vec{v} = 3\vec{e}_1 + 2\vec{e}_2$

Q: What if we use different set of vectors $\{\vec{j}_1, \vec{j}_2\}$ as a basis vectors for $\mathbb{R}^2$, what would be the "coordinates" of $\vec{v}$ with respect to $\{\vec{j}_1, \vec{j}_2\}$?

of $\vec{v}$ with respect to $\{\vec{j}_1, \vec{j}_2\}$?

Since $\{\vec{j}_1, \vec{j}_2\}$ is a basis there exist unique numbers x_1 and x_2, such that $\vec{v} = x_1 \vec{j}_1 + x_2 \vec{j}_2$. These numbers $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ are called the coordinates of $\vec{v}$ in the basis $\mathcal{J} = \{j_1, j_2\}$.

Sometimes we write $(\vec{v})_{\mathcal{J}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Remark: Note that in basis $\mathcal{J}$, the coordinates of vectors $\vec{j}_1$ and $\vec{j}_2$ are precisely $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, that is $(j_1)_{\mathcal{J}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $(j_2)_{\mathcal{J}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The basis defines the meaning of coordinates $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

In general, if we have a subspace $V \subseteq \mathbb{R}^n$ and a basis of V given by $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_m\}$, then for any $\vec{x} \in V$, $\vec{x} = \sum_{i=1}^{m} c_i \vec{v}_i$ and

$$(\vec{x})_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}.$$

Q: How do we translate between coordinate systems?

Example: Assume now that $(\vec{j}_1)_{\mathcal{E}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and that $(\vec{j}_2)_{\mathcal{E}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Say $(\vec{v})_{\mathcal{J}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, how do we find $(\vec{v})_{\mathcal{E}}$ for $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$? Well, $\vec{v} = -\vec{j}_1 + 2\vec{j}_2$ but $(\vec{j}_1)_{\mathcal{E}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $(\vec{j}_2)_{\mathcal{E}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and thus but $(\vec{j}_1)_{\mathcal{E}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $(\vec{j}_2)_{\mathcal{E}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and thus

$$(\vec{v})_{\mathcal{E}} = -1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} (\vec{v})_{\mathcal{J}} = \begin{pmatrix} -4 \\ 1 \end{pmatrix}$$

Intuition: The matrix $A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$ is a linear transform. from $\mathbb{R}^2$ to $\mathbb{R}^2$ that maps basis vectors of $\mathcal{E}$, $\vec{e}_1$ and $\vec{e}_2$ to basis vectors of $\mathcal{J}$, $\vec{j}_1$ and $\vec{j}_2$ respectively.

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Generally, given $\vec{v} = x\vec{e}_1 + y\vec{e}_2$ $\left((v)_{\mathcal{E}} = \begin{pmatrix} x \\ y \end{pmatrix} \right)$, applying A, we get the same linear combination

$$A\vec{v} = x \cdot A\vec{e}_1 + y \cdot A\vec{e}_2 = x \cdot \vec{j}_1 + y \cdot \vec{j}_2$$

with respect to the basis $\mathcal{J}$, of the resulting vector $A\vec{v}$. In particular $A\vec{e}_1 = \vec{j}_1$ and $A\vec{e}_2 = \vec{j}_2$.

Remark: Transformation $A: \mathbb{R}^2 \to \mathbb{R}^2$, geometrically, transforms basis $\mathcal{E}$ to $\mathcal{J}$, but numerically it „translates" vectors written in basis $\mathcal{J}$ into vectors written in basis $\mathcal{E}$.

Q: How about the other way around. Say we know the coordinates in basis $\mathcal{E}$, how do we find them in basis $\mathcal{J}$? By similar arguments, we take the inverse of A:

$$\{\text{vectors in } \mathcal{J}\} \underset{A^{-1}}{\overset{A}{\rightleftarrows}} \{\text{vectors in } \mathcal{E}\}$$

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}^{-1} = \overset{A^{-1}}{\begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}}.$$ So, for example if $(v)_{\mathcal{E}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$,

we will have $$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{3} \\ \frac{1}{3} \end{pmatrix}.$$

A is called the change of basis matrix.

In general, for two basis $\mathcal{B}$ and $\mathcal{C}$ of $V \subseteq \mathbb{R}^n$ with $\dim(V) = m$

$$A = \begin{pmatrix} | & | & & | \\ (\bar{c}_1)_{\mathcal{B}} & (\bar{c}_2)_{\mathcal{B}} & \cdots & (\bar{c}_m)_{\mathcal{B}} \\ | & | & & | \end{pmatrix}, \quad A(v)_{\mathcal{C}} = (v)_{\mathcal{B}}.$$

And and inverse matrix does the opposite.

Matrix of a linear transformation

Vectors are not the only thing with coordinates. Consider again the rotation T of $\mathbb{R}^2$ by $90°$. We showed that this transformation is given by a matrix $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ where colums of M represent where vectors $\vec{e}_1$ and $\vec{e}_2$ go in terms of, again, $\vec{e}_1$ and $\vec{e}_2$, that is,

$$\vec{e}_1 \overset{M}{\longmapsto} 0 \cdot \vec{e}_1 + 1 \vec{e}_2$$
$$\vec{e}_2 \overset{M}{\longmapsto} -1 \cdot \vec{e}_1 + 0 \cdot e_2$$

So, we follow" $\vec{e}_1, \vec{e}_2$ in the first place, and record"

So, we „follow" $\vec{e}_1, \vec{e}_2$ in the first place, and „record" the outcome also in $\vec{e}_1, \vec{e}_2$ in the second place.

Q: How would we describe the same 90° rotation with $\mathcal{J}$ basis? Where do $\vec{j}_1$ and $\vec{j}_2$ go in terms of $\vec{j}_1$ and $\vec{j}_2$? (Why is not $\left(\begin{pmatrix}0\\1\end{pmatrix}_{\mathcal{J}} \begin{pmatrix}-1\\0\end{pmatrix}_{\mathcal{J}}\right)$ correct?)

What we will do is, first start with any vector $(\vec{v})_{\mathcal{J}}$ written in $\mathcal{J}$ basis, using the cange of basis matrix, we will get $(\vec{v})_{\mathcal{E}}$, then we apply M, and then again change the basis back to $\mathcal{J}$ with an inverse matrix to get what „happend" to $\vec{v}$ in terms of $\mathcal{J}$:

$$\underbrace{\begin{pmatrix}2&-1\\1&1\end{pmatrix}^{-1}\overbrace{\underbrace{\begin{pmatrix}0&-1\\1&0\end{pmatrix}}_{\substack{\text{rotation by } 90^\circ\\ \text{in } \mathcal{E} \text{ basis.}}}\underbrace{\begin{pmatrix}2&-1\\1&1\end{pmatrix}(v)_{\mathcal{J}}}_{(v)_{\mathcal{E}}}}^{\text{Rotated } \vec{v} \text{ in } \mathcal{E} \text{ basis}}}_{\text{Rotated } \vec{v} \text{ in } \mathcal{J} \text{ basis}}$$

Therfore the composition $\begin{pmatrix}2&-1\\1&1\end{pmatrix}^{-1}\begin{pmatrix}0&-1\\1&0\end{pmatrix}\begin{pmatrix}2&-1\\1&1\end{pmatrix} = A^{-1}MA$ gives the matrix of the rotation by 90° in basis $\mathcal{J}$. In this case $A^{-1}MA = \begin{pmatrix}1/3 & -2/3\\ 5/3 & -1/3\end{pmatrix}$. So, if we take a vector $\vec{v}$, write it in basis $\mathcal{J}$ and apply $\begin{pmatrix}1/3 & -2/3 \\ \ldots\end{pmatrix}$ we will get a vector $\vec{w}$

take a vector $\vec{v}$, write it in basis $\mathcal{J}$ and apply $\begin{pmatrix} 1/3 & -2/3 \\ 7/3 & -1/3 \end{pmatrix}$, we will get a vector $\vec{w}$ written in basis $\mathcal{J}$, which is a $90°$ rotation of $\vec{v}$.

Of course, the same can be done for any basis $\mathcal{B}$ and $\mathcal{C}$ and a linear transformation T.

Similar matrices

Consider two $n \times n$ matrices A and B. We say that A is similar to B if there exists an invertible matrix S such that

$$AS = SB, \quad \text{or} \quad B = S^{-1}AS.$$

Basis change matrix

Thus, two matrices are similar if they represent the same linear transformation with respect to different bases.

Why can all this be important? (Snippet to eigenvalues and eigenvectors)

Lets consider a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ in basis $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ given by a matrix $\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = A$.

Let see what happens to $\operatorname{span}(\vec{v})$ under this transformation for some vector $\vec{v}$. For „most" vectors $A\vec{v} \notin \operatorname{span}(\vec{v})$, so a vector moved outside of $\operatorname{span}(\vec{v})$ subspace. But, it can happen that $A\vec{v} \in \operatorname{span}(\vec{v})$, or that is $A\vec{v} = k \cdot \vec{v}$ for some special $\vec{v}$.

For example,

$$\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} \quad (k = 1)$$

Whats more, then for any vector $\vec{w} \in \operatorname{span}\left(\begin{pmatrix} 0 \\ 3 \end{pmatrix}\right)$

Whats more, then for any vector $\vec{w} \in \text{span}\left(\begin{pmatrix}0\\3\end{pmatrix}\right)$ we have $\vec{w} = s \cdot \begin{pmatrix}0\\3\end{pmatrix}$ and $A\vec{w} = A\left(s\begin{pmatrix}0\\3\end{pmatrix}\right) = sA\begin{pmatrix}0\\3\end{pmatrix} = s\begin{pmatrix}0\\3\end{pmatrix} = \vec{w}$

so, $\vec{w}$ is also stretched by same factor as $\begin{pmatrix}0\\3\end{pmatrix}$ (namely, 1)

Also, note that

$$\begin{pmatrix}2 & 0\\1 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\2\end{pmatrix} = 2\cdot\begin{pmatrix}1\\1\end{pmatrix} \in \text{span}\begin{pmatrix}1\\1\end{pmatrix} \quad (k=2)$$

So for any vector $\vec{u} \in \text{span}\begin{pmatrix}1\\1\end{pmatrix}$ we also have $A\vec{u} = 2\cdot\vec{u}$.

Any other vector $\vec{x}$ (outside of $\text{span}\begin{pmatrix}1\\1\end{pmatrix}$ and $\text{span}\begin{pmatrix}0\\3\end{pmatrix}$) will get rotated, that is, $A\vec{x} \neq k\vec{x}$ for any k.

Def: A **nonzero** vector v is called an **eigenvector** of $n\times n$ matrix A (or, equivalently, of lin. tran $T: \mathbb{R}^n \to \mathbb{R}^n$) if $A\vec{v} = k\vec{v}$ for some scalar $k \in \mathbb{R}$. k is called an **eigenvalue** associated to eigenvector $\vec{v}$.

Q: How can we find such $\vec{v}$ an k?

Lets rewrite $A\vec{v} = k\vec{v}$ as

$$A\vec{v} - k\vec{v} = 0$$

$$(A - k\cdot I_n)\vec{v} = 0$$

For our example, $n=2$, so we know that system $(A - kI_2)\vec{v} = 0$ has a non-zero solution if and only if $A - kI_2$ is not invertible, which if and only if $A - \kappa I_2$ is not invertible, which happens only if $\det(A - \kappa I_2) = 0$. So, in our example, we will have

$$A - \kappa \cdot I_2 = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} \kappa & 0 \\ 0 & \kappa \end{pmatrix} = \begin{pmatrix} 2-\kappa & 0 \\ 1 & 1-\kappa \end{pmatrix}$$

$$\det(A - \kappa \cdot I_2) = \det\begin{pmatrix} 2-\kappa & 0 \\ 1 & 1-\kappa \end{pmatrix} = (2-\kappa)(1-\kappa), \quad \kappa_1 = 1, \ \kappa_2 = 2.$$

So, only possible eigenvalues $\kappa_1 = 1$ and $\kappa_2 = 2$.

Now, find eigenvalues by solving linear system:

$$\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 1 \cdot \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{cases} 2x = x \\ x + y = y \end{cases} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ t \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{cases} 2x = 2x \\ x + y = 2y \end{cases} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ t \end{pmatrix} \text{ as expected.}$$

Remark: Transformation does not have to have eigenvectors. (Example?)

Q: What happens if my basis vectors happen to be eigenvectors?

Example: Say $\vec{e}_1$ is scaled by -1 and $\vec{e}_2$ is scaled by two. So, $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 0 \\ 2 \end{pmatrix}$. Then, notice that matrix of this transformation will be diagonal: $\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$. In general, every time we have a diagonal matrix, all the basis vectors are diagonal: $\begin{pmatrix} & \\ 0 & 2 \end{pmatrix}$. In general, ... if we have a diagonal matrix, all the basis vectors are eigenvectors, and diagonal entries are eigenvalues.

Diagonal matrices are much easier to work with, for example, $\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}^{200} = \begin{pmatrix} (-1)^{200} & 0 \\ 0 & 2^{200} \end{pmatrix}$, but

$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{200}$ is much harder to compute.

Now, if our matrix happens to have „enough" eigenvectors, we can change the basis so that this eigenvectors become the basis. Then our matrix is guaranteed to become diagonal in this basis!

In our example:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

So, if we want to compute $\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}^{100}$, we first change to eigenvector basis, compute $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{100} = \begin{pmatrix} 1 & 0 \\ 0 & 2^{100} \end{pmatrix}$ and change back.

Remark: Not all matrices have basis consisting of eigenvalues.