Lecture 15

Spans and annihilators

Definition 19.8: Let $\mathcal{W} = (w_1, \dots, w_r)$ be a list of vectors in $\mathbb{R}^n$.

(a) $\text{span}(\mathcal{W})$ is the set of all vectors $v \in \mathbb{R}^n$ that can be expressed as a linear combination of the list $\mathcal{W}$.

(b) $\text{ann}(\mathcal{W})$ is the set of all vectors $u \in \mathbb{R}^n$ such that $u.w_1 = \cdots = u.w_r = 0$.

Remark 19.9: The terminology in (a) is related in an obvious way to the terminology used earlier: the list $\mathcal{W}$ spans $\mathbb{R}^n$ if and only if *every* vector in $\mathbb{R}^n$ is a linear combination of $\mathcal{W}$, or in other words $\text{span}(\mathcal{W}) = \mathbb{R}^n$.

Solution sets are annihilators

The solution set of a homogeneous linear system is an annihilator.

Example 19.10: Put

$$W = \left\{ a = \begin{bmatrix} w & x & y & z \end{bmatrix}^T \mid \begin{array}{rl} w + 2x - 2y - 3z & = 0 \text{ and} \\ w - x + y - z & = 0 \end{array} \right\}$$

$$= \left\{ a = \begin{bmatrix} w & x & y & z \end{bmatrix}^T \mid \begin{array}{rl} a.\begin{bmatrix} 1 & 2 & -2 & -3 \end{bmatrix}^T & = 0 \text{ and} \\ a.\begin{bmatrix} \quad & \quad & \quad & \quad \end{bmatrix}^T & = 0 \end{array} \right\}$$

$$= \text{ann}(\begin{bmatrix} 1 & 2 & -2 & -3 \end{bmatrix}^T, \begin{bmatrix} \quad & \quad & \quad & \quad \end{bmatrix}^T).$$

Example 19.11: Similarly, put

$$W = \left\{ v = \begin{bmatrix} p & q & r & s & t \end{bmatrix}^T \in \mathbb{R}^5 \mid \begin{array}{rl} p + q + r & = 0, \\ q + r + s & = 0 \text{ and} \\ r + s + t & = 0 \end{array} \right\}$$

If we put

$$a_1 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \end{bmatrix}^T \qquad a_2 = \begin{bmatrix} \quad & \quad & \quad & \quad & \quad \end{bmatrix}^T \qquad a_3 = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \end{bmatrix}^T$$

then the defining equations $p + q + r = q + r + s = r + s + t = 0$ can be rewritten as $a_1.v = a_2.v = a_3.v = 0$. This means that $W = \text{ann}(\quad, \quad, \quad)$.

Kernels are annihilators

Definition 19.13: Let A be an $m \times n$ matrix (so if $v \in \mathbb{R}^n$ then $Av \in \mathbb{R}^m$). We put $\ker(A) = \{v \in \mathbb{R}^n \mid Av = 0\}$, and call this the *kernel* of A.

Proposition 19.14: $\ker(A)$ is the annihilator of the transposed rows of A. In more detail, if

$$A = \begin{bmatrix} u_1^T \\ \hline \vdots \\ \hline u_m^T \end{bmatrix}.$$

then $\ker(A) = \text{ann}(u_1, \dots, u_m)$.

Proof.
We observed in Section 3 that $Av = \begin{bmatrix} u_1.v & \cdots & u_m.v \end{bmatrix}^T$.
Thus:

$$\begin{aligned} v \in \ker(A) &\Leftrightarrow Av = 0 \\ &\Leftrightarrow \quad = \cdots = \quad = 0 \\ &\Leftrightarrow v \in \text{ann}(u_1, \dots, u_m). \end{aligned}$$

□

Kernels are annihilators — example

Example 19.15: Consider the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 30 & 20 & 10 \\ 100 & 100 & 100 \end{bmatrix}$,

so if $v = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ we have $Av = \begin{bmatrix} x + 2y + 3z \\ 30x + 20y + 10z \\ 100x + 100y + 100z \end{bmatrix}$.

Thus:

$$\ker(A) = \{v \mid Av = 0\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid \begin{bmatrix} x + 2y + 3z \\ 30x + 20y + 10z \\ 100x + 100y + 100z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

$$= \left\{ v = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid v . \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = v . \begin{bmatrix} 30 \\ 20 \\ 10 \end{bmatrix} = v . \begin{bmatrix} \\ \\ \end{bmatrix} = 0 \right\}$$

$$= \operatorname{ann}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 30 \\ 20 \\ 10 \end{bmatrix}, \quad \begin{bmatrix} \\ \\ \end{bmatrix} \right)$$

Disguised span example

Example 19.16: Let $V \subseteq \mathbb{R}^4$ be the set of all vectors of the form

$$v = \begin{bmatrix} 2p+5q+3r \\ 9p-4q+2r \\ 3p+3q+3r \\ 8p-4q-5r \end{bmatrix}$$

$$v = \begin{bmatrix} 2p+5q+3r \\ 9p-4q+2r \\ 3p+3q+3r \\ 8p-4q-5r \end{bmatrix}$$

for some $p, q, r \in \mathbb{R}$. We can rewrite this as

$$v = p \begin{bmatrix} 2 \\ 9 \\ 3 \\ 8 \end{bmatrix} + q \begin{bmatrix} 5 \\ -4 \\ 3 \\ -4 \end{bmatrix} + r \begin{bmatrix} \\ \\ \\ \end{bmatrix}.$$

Thus, if we put

$$a = \begin{bmatrix} 2 \\ 9 \\ 3 \\ 8 \end{bmatrix} \qquad b = \begin{bmatrix} 5 \\ -4 \\ 3 \\ -4 \end{bmatrix} \qquad c = \begin{bmatrix} \\ \\ \\ \end{bmatrix},$$

then the general form for elements of V is $v = pa + qb + rc$.
In other words, the elements of V are precisely the vectors that can be expressed as a linear combination of a, b and c.
This means that $V = \operatorname{span}(a, b, c)$.

Disguised span example

Example 19.17: Let W be the set of all vectors in $\mathbb{R}^5$ of the form

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} t_3 - t_2 \\ t_1 + t_3 \\ t_2 - t_1 \\ t_1 + t_2 \\ t_3 + t_2 \end{bmatrix}$$

for some $t_1, t_2, t_3 \in \mathbb{R}$. This general form can be rewritten as

$$w = t_1 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + t_3 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = t_1 a_1 + t_2 a_2 + t_3 a_3,$$

where

$$a_1 = \begin{bmatrix} 0 & 1 & -1 & 1 & 0 \end{bmatrix}^T \quad a_2 = \begin{bmatrix} -1 & 0 & 1 & 1 & 1 \end{bmatrix}^T \quad a_3 = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \end{bmatrix}^T.$$

Thus, we see that a vector $w \in \mathbb{R}^5$ lies in W iff it can be expressed as a linear combination of the vectors a_1, a_2 and a_3.
This means that $W = \operatorname{span}(a_1, a_2, a_3)$.

Images are spans

Definition 19.18: Let A be an $m \times n$ matrix (so if $t \in \mathbb{R}^n$ then $At \in \mathbb{R}^m$). We write $\operatorname{img}(A)$ for the set of vectors $w \in \mathbb{R}^m$ that can be expressed in the form $w = At$ for some $t \in \mathbb{R}^n$. We call this the *image* of the matrix A.

Proposition 19.19: The image of A is just the span of the columns of A. In other words, if

$$A = \left[\begin{array}{c|c|c} v_1 & \cdots & v_n \end{array} \right]$$

then $\operatorname{img}(A) = \operatorname{span}(v_1, \ldots, v_n)$.

Proof: Recall from Section 3 that

$$At = \left[\begin{array}{c|c|c} v_1 & \cdots & v_n \end{array} \right] \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} = t_1 v_1 + \cdots + t_n v_n.$$

Note here that each t_i is a scalar (the i'th entry in the vector t) whereas v_i is a vector (the i'th column in the matrix A). Thus At is the sum of the arbitrary scalars t_i multiplied by the given vectors v_i; in other words, it is an arbitrary linear combination of $v_1, \ldots, v_n$. The claim is clear from this.

Images are spans — example

Let A be the matrix below, so $A = [a_1|a_2|a_3]$.

$$\begin{bmatrix} 2 & 30 & 4444 \\ 22 & 30 & 444 \\ 222 & 30 & 44 \\ 2222 & 30 & 4 \end{bmatrix} \qquad a_1 = \begin{bmatrix} 2 \\ 22 \\ 222 \\ 2222 \end{bmatrix}, \; a_2 = \begin{bmatrix} 30 \\ 30 \\ 30 \\ 30 \end{bmatrix}, \; a_3 = \begin{bmatrix} 4444 \\ 444 \\ 44 \\ 4 \end{bmatrix}.$$

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 30x_2 + 4444x_3 \\ 22x_1 + 30x_2 + 444x_3 \\ 222x_1 + 30x_2 + 44x_3 \\ 2222x_1 + 30x_2 + 4x_4 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 22 \\ 222 \\ 2222 \end{bmatrix} + x_2 \begin{bmatrix} 30 \\ 30 \\ 30 \\ 30 \end{bmatrix} + x_3 \begin{bmatrix} 4444 \\ 444 \\ 44 \\ 4 \end{bmatrix}$$
$$= x_1 a_1 + x_2 a_2 + x_3 a_3.$$

So
$$\begin{aligned} y \in \operatorname{img}(A) &\Leftrightarrow y \text{ can be written as } Ax \text{ for some } x \in \mathbb{R}^3 \\ &\Leftrightarrow y = x_1 a_1 + x_2 a_2 + x_3 a_3 \text{ for some } x_1, x_2, x_3 \in \mathbb{R} \\ &\Leftrightarrow y \text{ is a linear combination of } a_1, a_2 \text{ and } a_3 \\ &\Leftrightarrow y \in \operatorname{span}(a_1, a_2, a_3). \end{aligned}$$

So $\operatorname{img}(A) = \operatorname{span}(a_1, a_2, a_3)$.

Span and annihilator example

$\operatorname{span}(w_1, \dots, w_r) = \{$ linear combinations of $w_1, \dots, w_r\}$;
$\operatorname{ann}(w_1, \dots, w_r) = \{v \mid v.w_1 = \cdots = v.w_r = 0\}$

Consider the plane P in $\mathbb{R}^3$ with equation $x + y + z = 0$. More formally:

$$P = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \mid x + y + z = 0 \right\}.$$

If we put $v = \begin{bmatrix} x & y & z \end{bmatrix}^T$ and $t = \begin{bmatrix} & & \end{bmatrix}^T$, then we have $v.t = x + y + z$. It follows that

$$P = \{v \in \mathbb{R}^3 \mid v.t = 0\} = \operatorname{ann}(t).$$

On the other hand, if $x + y + z = 0$ then $z = $ so

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ -x-y \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + y \begin{bmatrix} \\ \\ \end{bmatrix}.$$

Thus, if we put $u_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$ and $u_2 = \begin{bmatrix} & & \end{bmatrix}^T$ then

$P = \{x\,u_1 + y\,u_2 \mid x, y \in \mathbb{R}\} = \{$ linear combinations of u_1 and $u_2\} = \operatorname{span}(u_1, u_2)$.

Span and annihilator example

Put

$$V = \{\begin{bmatrix} w & x & y & z \end{bmatrix}^T \in \mathbb{R}^4 \mid w + 2x + 3y + 4z = 4w + 3x + 2y + z = 0\}.$$

If we put $a = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^T$ and $b = \begin{bmatrix} & & & \end{bmatrix}^T$ then

$$w{+}2x{+}3y{+}4z = a.\begin{bmatrix} w & x & y & z \end{bmatrix}^T \qquad 4w{+}3x{+}2y{+}z = b.\begin{bmatrix} w & x & y & z \end{bmatrix}^T$$

so we can describe V as $\operatorname{ann}(a, b)$.

On the other hand, suppose we have a vector $v = \begin{bmatrix} w & x & y & z \end{bmatrix}^T$ in V, so that

$$\begin{aligned} w + 2x + 3y + 4z &= 0 && \text{(A)} \\ 4w + 3x + 2y + z &= 0 && \text{(B)} \end{aligned}$$

If we subtract 4 times (A) from (B) and then divide by -15 we get equation (C) below. Similarly, if we subtract 4 times (B) from (A) and divide by -15 we get (D).

$$\begin{aligned} \tfrac{1}{3}x + \tfrac{2}{3}y + z &= 0 && \text{(C)} \\ w + \tfrac{2}{3}x + \tfrac{1}{3}y &= 0 && \text{(D)} \end{aligned}$$

Span and annihilator example

$$\begin{aligned} V &= \{\begin{bmatrix} w & x & y & z \end{bmatrix}^T \in \mathbb{R}^4 \mid w + 2x + 3y + 4z = 4w + 3x + 2y + z = 0\} \\ &= \{\begin{bmatrix} w & x & y & z \end{bmatrix}^T \mid w = -\tfrac{2}{3}x - \tfrac{1}{3}y, \quad z = -\tfrac{1}{3}x - \tfrac{2}{3}y\} \\ &= \left\{ \begin{bmatrix} -\tfrac{2}{3}x - \tfrac{1}{3}y \\ x \\ y \\ -\tfrac{1}{3}x - \tfrac{2}{3}y \end{bmatrix} \mid x, y \in \mathbb{R} \right\} = \left\{ x \begin{bmatrix} -2/3 \\ 1 \\ 0 \\ -1/3 \end{bmatrix} + y \begin{bmatrix} \\ \\ \\ \end{bmatrix} \mid x, y \in \mathbb{R} \right\} \end{aligned}$$

Thus, if we put

$$c = \begin{bmatrix} -\frac{2}{3} \\ 1 \\ 0 \\ -\frac{1}{3} \end{bmatrix} \qquad d = \begin{bmatrix} \phantom{-\frac{1}{3}} \\ \\ \\ \phantom{-\frac{2}{3}} \end{bmatrix}$$

then

$$V = \{xc + yd \mid x, y \in \mathbb{R}\} = \operatorname{span}(c, d).$$

Annihilators are subspaces

A subspace must contain 0, and be closed under addition and scalar multiplication.

Proposition 19.23: For any list $\mathcal{W} = (w_1, \dots, w_r)$ of vectors in $\mathbb{R}^n$, the set

$$\operatorname{ann}(\mathcal{W}) = \{x \in \mathbb{R}^n \mid x.w_1 = \cdots = x.w_r = 0\}$$

is a subspace of $\mathbb{R}^n$.

Proof.

(a) The zero vector clearly has $0.w_i = 0$ for all i, so $0 \in \operatorname{ann}(\mathcal{W})$.

(b) Suppose that $u, v \in \operatorname{ann}(\mathcal{W})$. This means that $u.w_i = 0$ for all i, and that $v.w_i = 0$ for all i. It follows that $(u+v).w_i = u.w_i + v.w_i = 0 + 0 = 0$ for all i, so $u + v \in \operatorname{ann}(\mathcal{W})$. Thus, $\operatorname{ann}(\mathcal{W})$ is closed under addition.

(c) Suppose instead that $u \in \operatorname{ann}(\mathcal{W})$ and $t \in \mathbb{R}$. As before, we have $u.w_i = 0$ for all i. It follows that $(tu).w_i = t(u.w_i) = 0$ for all i, so $tu \in \operatorname{ann}(\mathcal{W})$. Thus, $\operatorname{ann}(\mathcal{W})$ is closed under scalar multiplication.

□

Spans are subspaces

A subspace must contain 0, and be closed under addition and scalar multiplication.

Proposition 19.24: For any list $\mathcal{W} = (w_1, \dots, w_r)$ of vectors in $\mathbb{R}^n$, the set $\operatorname{span}(\mathcal{W})$ (of linear combinations of $\mathcal{W}$) is a subspace of $\mathbb{R}^n$.

Proof.

(a) The zero vector can be written as a linear combination $0 = 0w_1 + \cdots + 0w_r$, so $0 \in \operatorname{span}(\mathcal{W})$.

(b) Suppose that $u, v \in \operatorname{span}(\mathcal{W})$. This means that for some sequence of coefficients $\lambda_i \in \mathbb{R}$ we have $u = \sum_i \lambda_i w_i$, and for some sequence of coefficients μ_i we have $v = \sum_i \mu_i w_i$. If we put $\nu_i = \lambda_i + \mu_i$ we then have $u + v = \sum_i \nu_i w_i$. This expresses $u + v$ as a linear combination of $\mathcal{W}$, so $u + v \in \operatorname{span}(\mathcal{W})$. Thus, $\operatorname{span}(\mathcal{W})$ is closed under addition.

(c) Suppose instead that $u \in \operatorname{span}(\mathcal{W})$ and $t \in \mathbb{R}$. As before, we have $u = \sum_i \lambda_i w_i$ for some sequence of coefficients λ_i. If we put $\kappa_i = t\lambda_i$ we find that $tu = \sum_i \kappa_i w_i$, which expresses tu as a linear combination of $\mathcal{W}$, so $tu \in \operatorname{span}(\mathcal{W})$. Thus, $\operatorname{span}(\mathcal{W})$ is closed under scalar multiplication.

□

Bases for subspaces

Definition 20.1: Let V be a subspace of $\mathbb{R}^n$. A *basis* for V is a linearly independent list $\mathcal{V} = (v_1, \dots, v_r)$ of vectors in $\mathbb{R}^n$ such that $\operatorname{span}(\mathcal{V}) = V$.

Definition 20.2: Let V be a subspace of $\mathbb{R}^n$. The *dimension* of V (written $\dim(V)$) is the maximum possible length of any linearly independent list in $\mathcal{V}$.

The empty list always counts as linearly independent, so $\dim(V) \geq 0$.
Any linearly independent list in $\mathbb{R}^n$ has length at most n, so $\dim(V) \leq n$.

Proposition 20.3: Let V be a subspace of $\mathbb{R}^n$, and put $d = \dim(V)$. Then any linearly independent list of length d in V is automatically a basis. In particular, V has a basis.

Independent lists of the right length are bases

Proposition: Let V be a subspace of $\mathbb{R}^n$, and put $d = \dim(V)$. Then any linearly independent list $\mathcal{V} = (v_1, \cdots, v_d)$ of length d in V is a basis.

Proof.
Let u be an arbitrary vector in V. Consider the list $\mathcal{V}' = (v_1, \dots, v_d, u)$. This is a list in V of length $d+1$, but d is the maximum possible length for any linearly independent list in V, so the list $\mathcal{V}'$ must be dependent. This means that there is a nontrivial relation

$$\lambda_1 v_1 + \cdots + \lambda_d v_d + \mu u = 0.$$

We claim that μ cannot be zero. Indeed, if $\mu = 0$ then the relation would become $\lambda_1 v_1 + \cdots + \lambda_d v_d = 0$, but $\mathcal{V}$ is linearly independent so this would give $\lambda_1 = \cdots = \lambda_d = 0$ as well as $\mu = 0$, so the original relation would be trivial, contrary to our assumption. Thus $\mu \neq 0$, so the relation can be rearranged as

$$u = -\frac{\lambda_1}{\mu} v_1 - \cdots - \frac{\lambda_d}{\mu} v_d,$$

which expresses u as a linear combination of $\mathcal{V}$. This shows that an arbitrary vector $u \in V$ can be expressed as a linear combination of $\mathcal{V}$, or in other words $V = \operatorname{span}(\mathcal{V})$. As $\mathcal{V}$ is also linearly independent, it is a basis for V. □

Any d-dimensional subspace is $\mathbb{R}^d$ in disguise

Proposition 20.4: Let V be a subspace of $\mathbb{R}^n$, and let $\mathcal{V} = (v_1, \dots, v_d)$ be a basis for V.
Define a function $\phi\colon \mathbb{R}^d \to V$ by $\phi(\lambda) = \lambda_1 v_1 + \cdots + \lambda_d v_d$.
Then there is an inverse function $\psi\colon V \to \mathbb{R}^d$ with $\phi(\psi(v)) = v$ for all $v \in V$, and $\psi(\phi(\lambda)) = \lambda$ for all $\lambda \in \mathbb{R}^d$. Moreover, both ϕ and ψ respect addition and scalar multiplication:

$$\begin{aligned} \phi(\lambda + \mu) &= \phi(\lambda) + \phi(\mu) & \qquad \phi(t\lambda) &= t\phi(\lambda) \\ \psi(v + w) &= \psi(v) + \psi(w) & \qquad \psi(tv) &= t\psi(v). \end{aligned}$$

Proof.
By assumption the list $\mathcal{V}$ is linearly independent and $\operatorname{span}(\mathcal{V}) = V$. Consider an arbitrary vector $u \in V$. As $u \in \operatorname{span}(\mathcal{V})$ we can write u as a linear combination $u = \sum_i \lambda_i v_i$ say, which means that $u = \phi(\lambda)$ for some λ. We claim that this λ is unique. Indeed, if we also have $u = \phi(\mu) = \sum_i \mu_i v_i$ then we can subtract to get $\sum_i (\lambda_i - \mu_i) v_i = 0$. This is a linear relation on the list $\mathcal{V}$, but $\mathcal{V}$ is assumed to be independent, so it must be the trivial relation. This means that all the coefficients $\lambda_i - \mu_i$ are zero, so $\lambda = \mu$ as required. It is now meaningful to define $\psi(u)$ to be the unique vector λ with $\psi(\lambda) = u$. Properties are left as an exercise. $\square$