Complete solution for a non-homogeneous system

Recall the system $A\mathbf{x} = \mathbf{b}$:

$$A = \begin{bmatrix} 1 & 3 & 2 & 10 & -7 \\ 2 & 6 & 6 & 28 & -20 \\ 1 & 3 & 5 & 22 & -16 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ -4 \end{bmatrix}$$

with rref form $R\mathbf{x} = \mathbf{d}$:

$$R = \begin{bmatrix} 1 & 3 & 0 & 2 & -1 \\ 0 & 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{d} = \begin{bmatrix} 11 \\ -3 \\ 0 \end{bmatrix}$$

Strang writes the general solution to $A\mathbf{x} = \mathbf{b}$ as:

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_p + \mathbf{x}_n \\ &= \begin{bmatrix} 11 \\ 0 \\ -3 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + u \begin{bmatrix} -2 \\ 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} + v \begin{bmatrix} 1 \\ 0 \\ 3 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

What's going on here? All the solutions for $A\mathbf{x} = \mathbf{b}$ are one <u>particular</u> solution to the (non-homogeneous) system plus any solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$, and the latter can be built out of the "special solutions" for the free variables:

$$\left[\begin{array}{ccccc|c} 1 & 3 & 0 & 2 & -1 & 11 \\ 0 & 0 & 1 & 4 & -3 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{array}{rcccc} \text{Soln to}: & A\mathbf{x} = \mathbf{b} & A\mathbf{x} = \mathbf{0} & A\mathbf{x} = \mathbf{0} & A\mathbf{x} = \mathbf{0} \\ y: & 0 & 1 & 0 & 0 \\ u: & 0 & 0 & 1 & 0 \\ v: & 0 & 0 & 0 & 1 \\ : & \begin{bmatrix} 11 \\ 0 \\ -3 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} -2 \\ 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \\ 3 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

Proof that every solution to $A\mathbf{x} = \mathbf{b}$ is a particular one plus a solution to $A\mathbf{x} = \mathbf{0}$:

If $\mathbf{x}_1, \mathbf{x}_2$ are both solutions, then

$$A(\mathbf{x}_2 - \mathbf{x}_1) = A\mathbf{x}_2 - A\mathbf{x}_1 = \mathbf{b} - \mathbf{b} = \mathbf{0} \ ,$$

so $\mathbf{x}_2 - \mathbf{x}_1 \in \mathbf{N}(A)$ and

$$\mathbf{x}_2 = \mathbf{x}_1 + (\mathbf{x}_2 - \mathbf{x}_1) \ .$$

Vector space proof of Rank+Nullity

Let T be a linear transformation from one fin dim vector space, V, to another, W. Let $r = \dim(T(V))$ and $k = \dim(N(T))$. Need to show $r + k = \dim(V)$.

Take bases $T(v_1), \ldots, T(v_r)$ of $T(V)$ (where $v_i \in V$) and $u_1, \ldots, u_k$ of $N(T)$. Claim $B = \{v_1, \ldots, v_r, u_1, \ldots, u_k\}$ is a basis for V, which will finish the proof.

Lin ind: Suppose

$$c_1 v_1 + \cdots + c_r v_r + d_1 u_1 + \cdots + d_k u_k = 0_V \ .$$

Apply T to both sides:

$$\begin{aligned} c_1 T(v_1) + \cdots + c_r T(v_r) + d_1 T(u_1) + \cdots + d_k T(u_k) &= T(0_V) \\ c_1 T(v_1) + \cdots + c_r T(v_r) &= 0_W \ . \end{aligned}$$

Because $T(v_i)$'s are lin ind, c_i's are all 0. So

$$d_1 u_1 + \cdots + d_k u_k = 0_V \ .$$

Because u_i's are lin ind, d_i's are also 0.

Span V: Take any v in V. Then $T(v)$ is in span of $T(v_1), \ldots, T(v_r)$, say

$$T(v) = c_1 T(v_1) + \cdots + c_r T(v_r) \ .$$

Now

$$T(v - (c_1 v_1 + \cdots + c_r v_r)) = 0_W \ ,$$

so $v - (c_1 v_1 + \cdots + c_r v_r) \in N(T)$, so for some d_i's,

$$v - (c_1 v_1 + \cdots + c_r v_r) = d_1 u_1 + \cdots + d_k u_k \ ,$$

i.e.,

$$v = c_1 v_1 + \cdots + c_r v_r + d_1 u_1 + \cdots + d_k u_k \ .$$

So B spans V. QED


I-A.I 2011

Challenge: Find a matrix with $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ in the column space and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ in the nullspace.

Answer: Impossible. The nullspace vectors are in $\mathbb{R}^3$, so $n = 3$. The column space vectors are lin ind, as are the nullspace vectors, so $r \geq 2$ and $n - r \geq 2$. This cannot be!

Challenge: Find a matrix with $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ in the column space and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in the nullspace.

Answer: Now we can do it. Let's use the column vectors as the first 2 columns, and build the 3rd column to make the nullspace vector work: 3rd column entry = negative of 1st and 2nd column entries:

$$\begin{bmatrix} 1 & 4 & -5 \\ 2 & 5 & -7 \\ 3 & 6 & -9 \end{bmatrix}.$$

Note 3rd is in the span of the first two. The rank must only be 2, because the nullspace has dimension at least 1.

The Other Two Subspaces

Definition

Let A be an $m \times n$ matrix. Then the *row space* of A is $\mathbf{R}(A) = \mathbf{C}(A^T)$, a subspace of $\mathbb{R}^n$; and the *left nullspace* of A is $\mathbf{N}(A^T)$, a subspace of $\mathbb{R}^m$.

Suppose $\mathrm{rank}(A) = r$:

- $\dim(\mathbf{C}(A)) = r$
- $\dim(\mathbf{N}(A)) = n - r$
- $\dim(\mathbf{R}(A)) = \dim(\mathbf{C}(A^T)) = \mathrm{rank}(A^T) = \mathrm{rank}(A) = r$
- $\dim(\mathbf{N}(A^T)) = m - r$

Orthogonal complement

Definition

Let S be a set of vectors in $\mathbb{R}^n$. Then the *orthogonal complement* of S, denoted $S^\perp$, is the set of all vectors that are orthogonal (perpendicular) to every vector in S.

You can check that:

- $S^\perp$ is a subspace of $\mathbb{R}^n$ (even if S isn't).
- S and $S^\perp$ have at most the zero vector in common.

Suppose S is a set of vectors in $\mathbb{R}^n$. Suppose we choose from S a basis for the span of S and use their transposes as the rows in an $r \times n$ matrix A — r because the rows are ind, so the number of rows is also the rank. Then $S^\perp = \mathbf{N}(A)$.

So $\dim(S^\perp) = n - r$.
So $\dim((S^\perp)^\perp) = n - (n - r) = r$.
So $(S^\perp)^\perp$ is the span of S — in fact, if S is a subspace, it <u>is</u> S.

We'll see shortly that, because the span of S and $S^\perp$ have only the zero vector in common, every vector in $\mathbb{R}^n$ is the sum of one from the span of S and one from $S^\perp$.

Orthogonalities in the four subspaces

Proposition

Let A be a matrix. Then $\mathbf{N}(A^T) = (\mathbf{C}(A))^{\perp}$.

Proof.

Suppose A is $m \times n$. For a vector $\mathbf{x}$ in $\mathbb{R}^m$, TFAE:

$$\begin{aligned}
\mathbf{x} \in (\mathbf{C}(A))^{\perp} &\iff \mathbf{x} \cdot (A\mathbf{y}) = 0 \text{ for all } \mathbf{y} \text{ in } \mathbb{R}^n \\
&\iff \mathbf{x}^T A\mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } \mathbb{R}^n \\
&\iff (A^T\mathbf{x})^T \mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } \mathbb{R}^n \\
&\iff (A^T\mathbf{x}) \cdot \mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } \mathbb{R}^n \\
&\iff A^T\mathbf{x} = \mathbf{0} \text{ in } \mathbb{R}^n \\
&\iff \mathbf{x} \in \mathbf{N}(A^T)
\end{aligned}$$

$\square$

Corollary

$(\mathbf{R}(A))^{\perp} = (\mathbf{C}(A^T))^{\perp} = \mathbf{N}(A).$

Proof.

(1) Replace A with A^T in the proposition.

(2) To say that a vector is a solution to $A\mathbf{x} = \mathbf{0}$ is to say that it is orthogonal to all the rows of A. $\square$

Challenge: Find a matrix with $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \\ -6 \end{bmatrix}$ in the row space and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in the nullspace.

Answer: Impossible: The nullspace vector isn't orthogonal to the second row space vector.

Chagrined subchallenge: So what vectors could be in the nullspace?

Answer: Any solution to the homogeneous system

$$\left[\begin{array}{cccc} 1 & 2 & -3 & 0 \\ 4 & 5 & -6 & 0 \end{array}\right].$$