Linear Maps, Matrices and Linear Sytems

Let $T : U \to V$ be linear. We defined $\ker(T) = \{x | x \in U, T(x) = 0\}$ and $im(T) = \{y | T(x) = y$ for some $x \in U\}$. According to the dimension equality we have that $\dim(\ker(T)) + \dim(im(T)) = \dim(U) = n$. We have that the linear map T is one-one or injective iff $\ker(T) = \{0\}$. If $U = V$ then T is injective iff T is surjective (onto). This is an easy consequence of the dimension equality.

For any linear map T we have a matrix representation. The matrix A of T depends on the chosen bases α and β of U and V, repectively:

$$Mat(T; \alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots \beta_m) = (a_{ij}) \text{ where } T(\alpha_j) = \Sigma_{i=1}^m a_{ij}\beta_i, j = 1, \ldots n; i = 1, \ldots m$$

The matrix for T is an $m \times n$ matrix where $\dim(V) = m, \dim(U) = n$. Each of the n columns of A contain the m components of $T(\alpha_j)$ with respect to the basis β_j.

Let x be a vector in U. If

$$x = \Sigma_{j=1}^n x_j\alpha_j, \text{then } T(x) = T(\Sigma_{j=1}^n x_j\alpha_j) = \Sigma_{j=1}^n x_j T(\alpha_j) = \Sigma_{j=1}^n x_j \Sigma_{i=1}^m a_{ij}\beta_i = \Sigma_{i=1}^m(\Sigma_{j=1}^n a_{ij}x_j)\beta_i = \Sigma_{j=1}^m y_i\beta_i$$

where

$$y_i = \Sigma_{j=1}^n a_{ij}x_j$$

Thus
$$\begin{pmatrix} a_{11} & a_{12} & \ldots & \ldots & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & \ldots & \ldots & a_{2n} \\ \ldots & \ldots & \ddots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ddots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & \ldots & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ \ldots \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ \ldots \\ y_m \end{pmatrix}$$

Thus the map $T(x) = y$ has a coordinate representation as $Ax = y$. with this in mind, we see that

$$\dim(im(T)) = \dim(\text{columnspace } A) = s, \dim(\ker(T)) = \dim(\text{solution space for } Ax = 0) = n - r$$

where r is the row rank of A. By the dimension equality we have that $s + (n - r) = n$ which is $s = r$ that is **row rank=column rank**.

Example 1:

Let $U = \mathbb{R}^3, V = \mathbb{R}^4$ then the matrix $A = \begin{pmatrix} 2 & 8 & -3 \\ 1 & 4 & 1 \\ -5 & 2 & 2 \\ 1 & -3 & 8 \end{pmatrix}$ stands for the linear map T

where $T\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}2\\1\\-5\\1\end{pmatrix}, T\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}8\\4\\2\\-3\end{pmatrix}, T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}-3\\1\\2\\8\end{pmatrix}$

where we have chosen the unit vectors as bases for $U = \mathbb{R}^3$ and $V = \mathbb{R}^4$.
We have that

$\begin{pmatrix}2 & 8 & -3\\1 & 4 & 1\\-5 & 2 & 2\\1 & -3 & 8\end{pmatrix}$ has rank 3. This means that $\ker(T) = \{0\}$, the map is injective and maps the 3 unit
vectors to three linearly independent vectors. The image of the vector $x \in \mathbb{R}^3$ is a vector in $\mathbb{R}^4$:

$\begin{pmatrix}x\\y\\z\end{pmatrix} \mapsto \begin{pmatrix}2x+8y-3z\\x+4y+z\\-5x+2y+2z\\x-3y+8z\end{pmatrix}$ $\ker(T)$ is the solution space for the homogeneous system which as we already saw consists only of the zero-vector of $\mathbb{R}^3$.
The map $T : \mathbb{R}^3 \to \mathbb{R}^4$ is injective. We can find a linear map $S : \mathbb{R}^4 \to \mathbb{R}^3$ such that $S \circ T : \mathbb{R}^3 \twoheadrightarrow \mathbb{R}^3$ is the identity. Because $T(e_1), T(e_2), T(e_3)$ are linearly independent, we can find a vector β_4 such that $\{\beta_1 = T(e_1), \beta_2 = T(e_2), \beta_3 = T(e_3), \beta_4\}$ form a basis of $\mathbb{R}^4$. We define S on this basis by $\beta_1 \mapsto e_1, \beta_2 \mapsto e_2, \beta_3 \mapsto e_3, \beta_4 \mapsto \alpha$ where α is any vector in $\mathbb{R}^3$. For example $\alpha = 0 \in \mathbb{R}^3$ is fine. Then $S(T(e_1)) = S(\beta_1) = e_1$ and similarly for the other unit vectors of $\mathbb{R}^3$. That is $S(T(e_i)) = e_i$. That is, the composition $S \circ T$ is the identity on the unit vectors of $\mathbb{R}^3$ and therefore the identity on $\mathbb{R}^3$. Because of $S \circ T = id_{\mathbb{R}^3}$ we have for general reasons that S is surjective and T injective. Something we showed for arbitrary maps. We also see that S is not uniquely determined by T. First β_4 is not unique and if we have found some β_4 we can assign any vector α in $\mathbb{R}^3$ as its image under S. The map S is unique only on $im(S)$ by assigning to $T(\alpha)$, the vector α.

Example 2. This is a simple example which makes the logic quite transparent. Let $T : \mathbb{R}^2 \twoheadrightarrow \mathbb{R}^3, e_1^2 \mapsto e_1^3, e_2^2 \mapsto e_2^3$. The matrix of T is $\begin{pmatrix}1 & 0\\0 & 1\\0 & 0\end{pmatrix}$ We add the third unit vector e_3^3 to $T(e_1^2) = e_1^3, T(e_2^2) = e_2^3$ and define the map S as
$S(T(e_1^2)) = e_1^2, S(T(e_2^2)) = e_2^3, S(e_3^3) = \alpha = \begin{pmatrix}a\\b\end{pmatrix}$, a, b arbitrarily chosen. Then

$$Mat(S; e_1^3, e_2^3, e_3^3; e_1^2, e_2^2) = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix}$$

and

$$Mat(S \circ T; e_1^2, e_2^2; e_1^2, e_2^2) = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This is an example where the product of two non-square matrices is square and invertible. Notice that

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{pmatrix}$ is also square but not the identity.

Matrices and linear maps can be identified. If A is an $m \times n$ –matrix, then $L_A : \mathbb{R}^n \twoheadrightarrow \mathbb{R}^m, X \mapsto AX = Y$ is linear and $Mat(L_A) = A$. This is because $Ae_j^n = A_j$ where A_j is the j^{th} –column vector of A.

Example 3.

The matrix

$\begin{pmatrix} 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix}$ stands for a map, A, from $\mathbb{R}^3$ into $\mathbb{R}^2$. The map is onto (why?), and therefore $\ker(A) + 2 = 3$, which gives us a one-dimensional null-space. How can we compute the kernel, that is find a basis? We have for

$\begin{pmatrix} 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix}$ as row echelon form: $\begin{pmatrix} 1 & 0 & 11 \\ 0 & 1 & -7 \end{pmatrix}$ This is $x = -11z, y = 7z$ or

$\ker(A)$ is the span of the vector $\begin{pmatrix} -11 \\ 7 \\ 1 \end{pmatrix}$

We have that the matrix of the composition of maps corresponds to the product of the matrices

If $S : U \to V, T : V \to W, A = Mat(S; \alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots \beta_m), B = Mat(T; \beta_1, \beta_2, \ldots \beta_m; \gamma_1, \gamma_2, \ldots, \gamma_l)$

then

$$Mat(T \circ S; \alpha_1, \alpha_2, \ldots, \alpha_n; \gamma_1, \gamma_2, \ldots, \gamma_l) = BA$$

In particular, $Mat(L_{BA}) = BA : L_{BA}(X) = (BA)X = B(AX) = L_B(L_A(X).$ Thus $L_{BA} = L_B \circ L_A$ and therefore $Mat(L_{BA}) = Mat(L_B \circ L_A)\_ = Mat(L_B)Mat(L_A) = BA$

The linear map $T : U \to V$ is invertible if there is a linear map $S : V \to U$ such that $S \circ T = id_U$ and $T \circ S = id_V$. For a linear map $T : U \to V$ to have an inverse, T^{-1}, it is necessary that $\dim U = \dim V$. (Proof?)

If $A = Mat(T; \alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots \beta_n)$ and $B = Mat(T^{-1}; \beta_1, \beta_2, \ldots \beta_n; \alpha_1, \alpha_2, \ldots, \alpha_n)$ then $B = A^{-1}$ where $AA^{-1} = A^{-1}A = I_n$, the identity $n \times n$ –matrix.

We have the following important result:

For an $n \times n$ –matrix the following are equivalent:

$Ax = 0$ has only the trivial solution;

$Ax = y$ has for every y exactly one solution x

A has an inverse.

All of this follows from the theorem that a linear map on a finite dimensional vector space is injective if an only if it is surjective.

Now, how can we find the inverse of a matrix? While the book postpones this up to a later chapter, see p.100, Example 2, using our current knowledge on linear maps this is actually quite trivial to do. Let us explain this on that example:

$A = \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$ stands for the linear map $T = L_A : \mathbb{R}^2 \to \mathbb{R}^{2,}$ where

$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 7 \\ 3 \end{pmatrix}$. That is $A = Mat(T; e_1, e_2; e_1, e_2)$ where the e_i are the unit vectors in $\mathbb{R}^2$. We have

$T^{-1}\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $T^{-1}\begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then obviously,

$Mat(T^{-1}; Te_1, Te_2; e_1, e_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and this is not what we want. We want

$Mat(T^{-1}; e_1, e_2; e_1, e_2) = A^{-1}$. For this we need to find

$T^{-1}(e_1) = \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix}, T^{-1}(e_2) = \begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix}$ But for this we need to express the unit vectors as linear combinations of $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$ and

$\begin{pmatrix} 7 \\ 3 \end{pmatrix} : x\begin{pmatrix} 5 \\ 2 \end{pmatrix} + y\begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ This is the inhomogeneous linear system of

2 equations in 2 unknowns with augmented matrix

$\begin{pmatrix} 5 & 7 & 1 \\ 2 & 3 & 0 \end{pmatrix}$ which has the row echelon form: $\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \end{pmatrix}$. Hence:

$x = 3, y = -2$. And from:

$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3\begin{pmatrix} 5 \\ 2 \end{pmatrix} - 2\begin{pmatrix} 7 \\ 3 \end{pmatrix}$ we get

$T^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3T^{-1}\begin{pmatrix} 5 \\ 2 \end{pmatrix} - 2T^{-1}\begin{pmatrix} 7 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 2\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$

and similarly for the second column of the inverse. Actually we can work out simultaneously both inhomogeneous systems where the right hand sides are the unit vectors

$\begin{pmatrix} 5 & 7 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{pmatrix}$, row echelon form: $\begin{pmatrix} 1 & 0 & 3 & -7 \\ 0 & 1 & -2 & 5 \end{pmatrix}$

Thus, $Mat(T^{-1}; e_1, e_2; e_1, e_2) = \begin{pmatrix} 3 & -7 \\ -2 & 5 \end{pmatrix}$

Let A be an $n \times n$ –matrix where L_A has an inverse. Then A has an inverse A^{-1}. Then if A_j is the j^{th} –column of A then $L_A(e_j) = A_j$. Then if

$$x_{1j}A_1 + x_{2j}A_2 + \ldots + x_{nj}A_n = e_j, \text{ where } e_j \text{ is the } j-\text{th unit vector}$$

then applying L_A^{-1} to this equation gives:

$$x_{1j}e_1 + x_{2j}e_2 + \ldots + x_{nj}en = L_A^{-1}(e_j)$$

This is $L_A^{-1}(e_j) = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} = X_j$ and the matrix with columns X_j is the matrix of L_A^{-1} and the inverse of A.

Let $(A \mid I_n)$ be the matrix A augmented by the n –columns of unit vectors $e_1, e_2, \ldots, e_n$. Then using the elementary row operations transforms A into I_n and I_n into A^{-1}.

$$(A \mid I_n) \overset{\text{elementary row operations}}{\Longrightarrow} (I_n \mid A^{-1})$$

Actually, $AA^{-1} = I_n$. This tells us that $L_{A^{-1}}$ is injective. But then it must be also surjective, that is AA^{-1} also.