

ÁLGEBRA LINEAL, PARTE 4: MAS TRANSFORMACIONES LINEALES¹

We continue following the relevant sections from our main references:

- ch. 5-7, 12 of Mota, Rumbos *Álgebra lineal en $\mathbb{R}^n$* ,
- ch. 8,5 of Vargas, Wachtel *Álgebra lineal I*.

Our basic theme, a linear equation

$$(*) \quad A\mathbf{x} = \mathbf{b}$$

involves two vectors $\mathbf{x}, \mathbf{b}$ in *non-symmetric roles*. More precisely:

- If we are given $\mathbf{x}$ and ask ¿what is $\mathbf{b}$? our answer is immediate: we just carry out the matrix multiplication of A with $\mathbf{x}$ to have $\mathbf{b}$.
- If we are given $\mathbf{b}$ and ask ¿what is $\mathbf{x}$? our answer is less immediate: we will need to study the system deciding first of all if it is consistent and if so to then to describe its solutions (questions we may settle using e.g. Gauss-Jordan elimination).

In this note we will consider those circumstances (when A is ‘invertible’) in which we may reverse the roles of $\mathbf{x}, \mathbf{b}$ to have $(*)$ equivalent to

$$(*)' \quad \mathbf{x} = A^{-1}\mathbf{b}$$

for an appropriate matrix A^{-1} (the ‘inverse matrix’ of A). For $(*)'$ our second question (given $\mathbf{b}$ and asking for $\mathbf{x}$) is now the simpler immediate case.

Contents

1	Invertible linear transformations	2
1.1	some examples	4
2	changing basis	8

¹Connor Jackman, connor.jackman@itam.mx, 10.2026

1 Invertible linear transformations

For a general function $f : X \rightarrow Y$ between two sets, recall:

- f is called *surjective* (or onto) when for any $y \in Y$ there is some $x \in X$ such that $f(x) = y$. That is when there always *exists* some solution to $f(x) = y$, or, in other words, when the image of f covers all of Y :

$$f \text{ surjective} \iff \text{im}(f) = Y$$

- f is called *injective* (or one-to-one) when for any $y \in Y$ there is at most one $x \in X$ such that $f(x) = y$. Equivalently: $f(x) = f(x') \iff x = x'$, or, in other words, when for given $y \in \text{im}(f) \subset Y$ there is a *unique solution* to $f(x) = y$.
- f is called *bijective* (or invertible) when it is surjective and injective. That is any $y \in Y$ corresponds via $f(x) = y$ to a unique $x \in X$. The *inverse* of a bijective map is $f^{-1} : Y \rightarrow X$ via $f^{-1}(y) = x \iff f(x) = y$. That is f^{-1} sends given $y \in Y$ to the unique solution $x \in X$ of $f(x) = y$. Determining the inverse to an invertible map thus amounts to finding the general solutions of $f(x) = y$ ('undoing' the action of applying f).

For linear transformations (or linear maps)

Proposition 1. A linear transformation $L : V \rightarrow W$ (between two finite dimensional vector spaces) is:

1. *surjective (onto)* iff $\text{im}(L) = W$ iff $\text{rank}(L) = \dim(W)$,
2. *injective (one-to-one)* iff $\ker(L) = \{\mathbf{0}\}$ iff $\text{null}(L) = 0$,
3. *bijective (invertible)* iff $\dim V = \dim W$ and L is onto iff $\dim V = \dim W$ and L is one-to-one.

Proof. The first item is the definition of surjective and rank (recall: if $U \subset W$ is a subspace with $\dim U = \dim W < \infty$ then $U = W$). For the 2nd item, if L is injective then $L(\mathbf{v}) = \mathbf{0} = L(\mathbf{0}) \implies \mathbf{v} = \mathbf{0}$ so that L injective implies its kernel is trivial. Conversely when $L(\mathbf{v}) = L(\mathbf{v}')$ we have (by linearity of L) that $L(\mathbf{v} - \mathbf{v}') = \mathbf{0} \implies \mathbf{v} - \mathbf{v}' \in \ker(L) = \{\mathbf{0}\} \implies \mathbf{v} - \mathbf{v}' = \mathbf{0} \implies \mathbf{v} = \mathbf{v}'$, ie L is one-to-one. The 3rd item follows from the first two and rank-nullity. Namely if $\dim V = \dim W$ and L is onto (so $\text{rank}(L) = \dim W = \dim V$) then, by rank-nullity, we have $\dim V = \text{null}(L) + \text{rank}(L) = \text{null}(L) + \dim(V) \implies \text{null}(L) = 0$ so that L is also injective. Likewise if $\dim V = \dim W$ and L is injective then, by rank-nullity, we have $\dim W = \dim V = \text{null}(L) + \text{rank}(L) = \text{rank}(L)$ so that L is also surjective. $\square$

So, regarding invertible linear transformations, it suffices to consider when $\dim V = \dim W$ are of the same dimension: e.g. $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$, or, in terms of matrices those linear maps which correspond to *square matrices* ($n \times n$).

The last concepts have the following useful characterizations:

Proposition 2. *For $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear and invertible, then also its inverse $L^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is linear. In fact the inverse of L is the unique function satisfying*

- $L \circ L^{-1} = I_{n \times n}$, and, $L^{-1} \circ L = I_{n \times n}$ (for $I_{n \times n} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ the identity).

Proof. By definition, given $\mathbf{b} \in \mathbb{R}^n$ then $L^{-1}(\mathbf{b}) = \mathbf{x}$ is the unique $\mathbf{x} \in \mathbb{R}^n$ solving $L(\mathbf{x}) = \mathbf{b}$. The linearity of the inverse, L^{-1} , now follows from that of L : if $L(\mathbf{x}) = \mathbf{b}$ then $L(c\mathbf{x}) = c\mathbf{b}$, thus $L^{-1}(c\mathbf{b}) = c\mathbf{x} = cL^{-1}(\mathbf{b})$. Likewise for sums. The functional characterizations $LL^{-1} = I = L^{-1}L$ express that the inverse of L ‘undoes’ the operation of applying L . Indeed given $\mathbf{b}$ then $L^{-1}(\mathbf{b}) = \mathbf{x}$ is the unique solution to $L(\mathbf{x}) = \mathbf{b}$, ie $L(L^{-1}(\mathbf{b})) = \mathbf{b}$, ie $LL^{-1} = I$ is the identity (likewise for $L^{-1}L = I$). That L^{-1} is uniquely determined by these equations, suppose some $M : \mathbb{R}^n \rightarrow \mathbb{R}^n$ also satisfies $LM = I = ML$. Then:

$$M = I \circ M = (L^{-1} \circ L) \circ M = L^{-1} \circ (L \circ M) = L^{-1} \circ I = L^{-1}.$$

□

Definition 1. *For $L : V \rightarrow W$ we call $M : W \rightarrow V$ a right inverse of L when:*

$$L \circ M = I_W$$

for $I_W : W \rightarrow W, \mathbf{w} \mapsto \mathbf{w}$ the identity transformation (of W). Likewise we call $K : W \rightarrow V$ with $K \circ L = I_V$ a left inverse of L .

Proposition 3. *For a map $L : V \rightarrow W$ then*

1. *L has some right inverse iff L is surjective (onto).*
2. *L has some left inverse iff L is injective (one-to-one).*

Proof. If L has a right inverse M , then for any $\mathbf{w} \in W$ set $\mathbf{v} = M(\mathbf{w}) \in V$ so that: $\mathbf{w} = I(\mathbf{w}) = (L \circ M)(\mathbf{w}) = L(M(\mathbf{w})) = L(\mathbf{v})$, ie $im(L) = W$ so L is onto. On the other hand, if L is onto then for each $\mathbf{w} \in W$ there exists some $\mathbf{v} \in V$ so that $L(\mathbf{v}) = \mathbf{w}$. Choose one such $\mathbf{v}_{\mathbf{w}}$ for each $\mathbf{w}$ and take $M(\mathbf{w}) = \mathbf{v}_{\mathbf{w}}$ as a right inverse. Similarly if L has a left inverse K , suppose that $L(\mathbf{v}) = L(\mathbf{v}')$ and apply K to both sides to have $\mathbf{v} = I(\mathbf{v}) = K(L(\mathbf{v})) = K(L(\mathbf{v}')) = I(\mathbf{v}') = \mathbf{v}'$ (using $K \circ L = I$) so that L is 1-1. Conversely, if L is 1-1, then for each $\mathbf{w} \in im(L)$ there is a unique $\mathbf{v}_{\mathbf{w}} \in V$ with $L(\mathbf{v}_{\mathbf{w}}) = \mathbf{w}$. Define $K(\mathbf{w}) = \mathbf{v}_{\mathbf{w}}$ for $\mathbf{w} \in im(L)$, and for $\mathbf{w}' \in W \setminus im(L)$ choose any $\mathbf{v}_* \in V$ and take $K(\mathbf{w}') = \mathbf{v}_*$ to have a left inverse. □

EXERCISES.

- Suppose L is invertible, with inverse L^{-1} . Show also L^{-1} is invertible and $(L^{-1})^{-1} = L$ (suggestion: use prop. 2).
- Suppose $L, M : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are both invertible. Show that their composition $L \circ M$ is also invertible and $(L \circ M)^{-1} = M^{-1} \circ L^{-1}$.
- Given a linear map $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$. If L admits a left inverse, is the left inverse always unique? Same question for a right inverse.

1.1 some examples

- We first consider how we may use Gauss-Jordan elimination to compute the inverse of an invertible linear transformation.

So consider a linear transformation induced by a square $n \times n$ matrix A :

$$\mathbb{R}^n \rightarrow \mathbb{R}^n, \mathbf{x} \mapsto A\mathbf{x}.$$

Checking if A is invertible we know how to settle: we may reduce A to FE form and count its rank/nullity. Then A will be invertible iff $\text{rank}(A) = n$, $\text{null}(A) = 0$ (prop. 1). Now however we desire more. We would like to find an explicit formula for said inverse. From prop. 2 we know that this inverse will be also a linear transformation, say corresponding to an $n \times n$ matrix A^{-1} , which solves (recall: composition $\equiv$ matrix multiplication)

$$AA^{-1} = I_{n \times n}, \quad A^{-1}A = I_{n \times n}.$$

Let us consider the 1st equation, column by column. We will see it amounts to a series of questions we already know how to solve, namely a series of linear equations. Set:

$$\mathbf{c}_j = \text{col}_j(A^{-1}), \quad \mathbf{e}_j = \text{col}_j(I_{n \times n})$$

for the j 'th column of A^{-1} (to be determined) and the j 'th column of $I_{n \times n}$ (we know already $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ are the standard basis vectors for $\mathbb{R}^n$). Then to solve for the columns, $\mathbf{c}_j$ of A^{-1} we have from $AA^{-1} = I_{n \times n}$ the series of linear systems:

$$A\mathbf{c}_1 = \mathbf{e}_1, \quad A\mathbf{c}_2 = \mathbf{e}_2, \quad \dots, \quad A\mathbf{c}_n = \mathbf{e}_n.$$

To resolve them all simultaneously we form the augmented matrix:

$$[A \mid I_{n \times n}]$$

and reduce to FER form to find the desired columns $\mathbf{c}_j$ of A^{-1} . In fact, A being invertible, we will find always:

$$[A \mid I_{n \times n}] \sim [I_{n \times n} \mid A^{-1}]$$

upon completing the Gauss-Jordan algorithm.

For a concrete example consider

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 3 & 4 \end{pmatrix}.$$

To determine an inverse (if it exists) we form the augmented matrix, and reduce with Gauss-Jordan:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 & 1 & 0 \\ 2 & 3 & 4 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right)$$

and we have our inverse matrix:

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

Given some $\mathbf{b} \in \mathbb{R}^3$, to find the solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$ we just need to matrix multiply $\mathbf{x} = A^{-1}\mathbf{b}$.

- There is another common way to resolve the same question from the last example (given a matrix, to find –when it exists– its inverse matrix).

The method is the same (Gauss-Jordan), however the interpretation and presentation of the result (the inverse matrix) is different. Namely we will obtain the inverse matrix ‘factorized’ as a certain product of ‘elementary’ (ie simple) matrices.

Consider our algorithm presented in the last example. We form the augmented matrix $[A \mid I]$ and begin reducing towards FER form. Each step in the reduction consists in applying some elementary operation (swap some rows, re-scale a row, or add a multiple of one row to another row). Each of these operations can be itself effected by multiplying through the augmented matrix by a certain simple ‘elementary’ matrix.

Let M be any $m \times n$ matrix.

- to swap row j of M with row k of M take SM for S the $m \times m$ matrix with columns:

$$col_j(S) = \mathbf{e}_k, \quad col_k(S) = \mathbf{e}_j$$

and $col_l(S) = \mathbf{e}_l$ for $l \notin \{j, k\}$ (here $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis for $\mathbb{R}^m$).

- to re-scale row j of M by $c \neq 0$ take RM for R the $m \times m$ matrix with columns:

$$col_j(R) = c\mathbf{e}_j$$

and $col_l(R) = \mathbf{e}_l$ for $l \neq j$.

- to add a multiple c of row j of M to row k of M take QM for Q the $m \times m$ matrix with columns:

$$col_j(Q) = \mathbf{e}_j + c\mathbf{e}_k$$

and $col_l(Q) = \mathbf{e}_l$ for $l \neq j$.

In our present situation (A is a square $n \times n$ matrix) these matrices representing the elementary operations are easy to remember since they are just those matrices we obtain by applying said elementary operation to the identity matrix $I_{n \times n}$.

Thus, as an alternative to obtain the inverse matrix to A , we may avoid forming the augmented matrix and simply reduce A to FER form while keeping track of the elementary operations we apply along the way. The inverse of A will then be given by the product of the corresponding elementary matrices to these operations.

For a concrete example consider:

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}.$$

Then reducing to FER form, while recording our operations, we have for example:

$$\begin{aligned} \begin{pmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix} &\sim_{R_1 \leftrightarrow R_3} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \sim_{R_2 \rightarrow R_2 - R_1, R_3 \rightarrow R_3 - 3R_1} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & -1 & -5 \end{pmatrix} \\ &\sim_{R_3 \rightarrow R_3 + R_2} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -4 \end{pmatrix} \sim_{R_3 \rightarrow -\frac{1}{4}R_3} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \sim_{R_1 \rightarrow R_1 - R_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &\sim_{R_1 \rightarrow R_1 - R_3, R_2 \rightarrow R_2 - R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

which yield the inverse as the following product of elementary matrices:

$$A^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

- We can also use these methods to find –when they exist– right (or left) inverses to matrices that need not be square.

So, consider a linear transformation $\mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{x} \mapsto A\mathbf{x}$ induced by an $m \times n$ matrix A .

- Suppose we seek a right inverse to A , that is we seek an $n \times m$ matrix B such that

$$AB = I_{m \times m}.$$

We know already (prop. 3) this will be possible iff A is surjective, ie $\text{rank}(A) = m$. By rank-nullity we must have, at least that $n = \text{rank}(A) + \text{null}(A) \geq m$. We can in this case apply the same reasoning as we used to find an inverse and find the columns of B by reducing the augmented matrix: $[A \mid I_{m \times m}]$.

As a concrete example, consider

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 2 \end{pmatrix}$$

which represents a certain linear map $\mathbb{R}^4 \rightarrow \mathbb{R}^2$. To seek a right inverse, A must have rank two, which we can see is the case since eg its first two columns are not multiples (so are independent). Thus there is some right inverse, and we can find them by reducing:

$$\left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 1 & 0 \\ 1 & 1 & 2 & 2 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{cccc|cc} 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 & -1 \end{array} \right).$$

So, as a first column of a right inverse $B : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ we can take any $(x_1, x_2, x_3, x_4)^T \in \mathbb{R}^4$ such that:

$$x_1 + x_3 = -1, \quad x_2 + x_3 + 2x_4 = 1$$

and as a 2nd column of B any $(y_1, y_2, y_3, y_4)^T \in \mathbb{R}^4$ such that

$$y_1 + y_3 = 0, \quad y_2 + y_3 + 2y_4 = -1$$

that is upon labeling the free variables $a = x_3, b = y_3, c = x_4, d = y_4$ we have the right inverses:

$$B = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -a & -b \\ -a - 2c & -b - 2d \\ a & b \\ c & d \end{pmatrix}$$

for any choice of the parameters $a, b, c, d \in \mathbb{R}$.

- Suppose we seek a left inverse to A , that is we seek an $n \times m$ matrix B such that

$$BA = I_{n \times n}.$$

First note (prop. 3) that this will be possible iff A is injective iff $\ker A = \{0\}$, ie the nullity of A is zero. By rank-nullity we must have, at least, that $n = \text{rank}(A) \leq m$.

We can in this case apply the same reasoning from before by taking the transpose:

$$A^T B^T = I_{n \times n}$$

in order to, as before, solve for the columns of B^T , ie the rows of B by Gauss-Jordan reduction.

As a concrete example, consider

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 4 & 6 \end{pmatrix}$$

which represents a certain linear map $\mathbb{R}^2 \rightarrow \mathbb{R}^3$. To seek a left inverse, A must have rank two (nullity zero) which we can see is the case since its two columns are not multiples of each other. Thus there is some

left inverse, $B : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, and we can find them by reducing the transpose:

$$\left(\begin{array}{ccc|cc} 1 & 3 & 4 & 1 & 0 \\ 2 & 4 & 6 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|cc} 1 & 0 & 1 & -2 & \frac{3}{2} \\ 0 & 1 & 1 & 1 & -\frac{1}{2} \end{array} \right).$$

So, as a first column of B^T (first row of B) we can take any $(x_1, x_2, x_3)^T$ solving:

$$x_1 + x_3 = -2, \quad x_2 + x_3 = 1$$

and as a second column of B^T (second row of B) any $(y_1, y_2, y_3)^T$ solving

$$y_1 + y_3 = \frac{3}{2}, \quad y_2 + y_3 = -\frac{1}{2}$$

that is, upon labeling the free variables $a = x_3, b = y_3$ we have the left inverses:

$$B = \begin{pmatrix} -2 & 1 & 0 \\ \frac{3}{2} & -\frac{1}{2} & 0 \end{pmatrix} + \begin{pmatrix} -a & -a & a \\ -b & -b & b \end{pmatrix}$$

for any choice of the parameters $a, b \in \mathbb{R}$.

2 changing basis

A common situation in which we find invertible matrices is when we ‘change the basis’ in which we represent a given linear map. Changing basis is a certain type of coordinate change. In terms of an $m \times n$ matrix A we seek to simplify the form of A under the action:

$$A \mapsto XAY$$

for X an $m \times m$ invertible matrix and Y an $n \times n$ invertible matrix.

To explain how this comes about, let us recall:

Proposition 4. *Given a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V then any $\mathbf{x} \in V$ has a unique expansion:*

$$\mathbf{x} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n, \quad (x_j \in \mathbb{R})$$

in this basis. We call the bijection (an invertible linear map)

$$\mathbb{R}^n \rightarrow V, \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n$$

the coordinates on V associated to this basis.

Proof. Since $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis, it is (1) a linearly independent list, and, (2) spans V . Given any $\mathbf{x} \in V$ then by (2) (since $V = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$) we have $\mathbf{x} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$ for some $x_j \in \mathbb{R}$. To see these coefficients are unique, we note that if also $\mathbf{x} = x'_1\mathbf{v}_1 + \dots + x'_n\mathbf{v}_n = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$ then $\mathbf{0} = (x_1 - x'_1)\mathbf{v}_1 + \dots + (x_n - x'_n)\mathbf{v}_n$ which implies by (1) (that $\mathbf{v}_j$ are independent) that $x_1 - x'_1 = \dots = x_n - x'_n = 0$, ie $x_j = x'_j$ thus this representation is unique. $\square$

Upon choosing bases, any linear transformation has a certain associated matrix representation:

Proposition 5. *Let $L : V \rightarrow W$ be a linear transformation between an n -dimensional vector space V and an m -dimensional vector space W .*

Then to each choice of bases $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \dots, \mathbf{w}_m$ of W we have the associated $m \times n$ matrix representation of L in these bases as the $m \times n$ matrix whose j 'th column is

$$\text{col}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad (j = 1, \dots, n),$$

for $L(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \dots + a_{mj}\mathbf{w}_m$. This is the unique $m \times n$ matrix, A , so that the linear map:

$$P^{-1} \circ L \circ Q : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

is given by matrix multiplication by A . Here $Q : \mathbb{R}^n \rightarrow V$ and $P : \mathbb{R}^m \rightarrow W$ are the invertible linear maps (coordinates) associated to these choices of bases. The situation is usually summarized in the following commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ Q \uparrow & & \downarrow P^{-1} \\ \mathbb{R}^n & \xrightarrow{m.m.} & \mathbb{R}^m \end{array}$$

where the bottom dashed arrow is by matrix multiplication ($m.m.$).

Proof. For the invertible linear maps (coordinates) associated to our bases

$$Q : \mathbb{R}^n \rightarrow V, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n, \text{ and } P : \mathbb{R}^m \rightarrow W, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \mapsto y_1\mathbf{w}_1 + \dots + y_m\mathbf{w}_m$$

we have the linear map $P^{-1} \circ L \circ Q : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which sends an element of the standard basis $\mathbf{e}_j \in \mathbb{R}^n$ to:

$$\mathbf{e}_j \xrightarrow{Q} \mathbf{v}_j \xrightarrow{L} L(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \dots + a_{mj}\mathbf{w}_m \xrightarrow{P^{-1}} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} =: \text{col}_j$$

and so a general element $\mathbf{x} \in \mathbb{R}^n$ is sent by $P^{-1} \circ L \circ Q$ to

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1 \text{col}_1 + \dots + x_n \text{col}_n = A\mathbf{x}$$

for A the $m \times n$ matrix with columns $\text{col}_1, \dots, \text{col}_n$ as defined above. $\square$

EXAMPLES.

- Changing basis is a special case of changing coordinates. Transformations are a fundamental operation, and lead to the general technique of: given a (complicated) problem, can we transform it under some coordinate change into an equivalent simpler problem? For example consider the linear map $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ given in the standard basis by

$$A = \begin{pmatrix} 3 & -5 \\ 2 & -3 \end{pmatrix}$$

In a different basis: $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ this map will be represented by a different matrix. Namely consider the change of coordinates $Q : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\begin{pmatrix} X \\ Y \end{pmatrix} \mapsto X\mathbf{v}_1 + Y\mathbf{v}_2$, given in the standard bases $(\mathbf{e}_1, \mathbf{e}_2)$ by the matrix (we will still denote by ‘ Q ’)

$$Q = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

and with inverse $Q^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. Then, in this new basis, our linear map is represented by the matrix:

$$A' = Q^{-1}AQ = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Indeed, a simpler looking form (more zeroes, ones) than the original A !

Of course this example is at the moment slightly artificial since we have just produced a simplifying change of basis out of thin air. In order for this idea to be more convincing we require some method in order to find those bases in which the form of a given linear map (matrix) will be ‘simplified’ (this will be part of our next topic: Eigenspaces).

- In the last example we are looking at a rather special type of change of basis, a linear map $A : V \rightarrow V$ where we use the *same* basis on the domain and co-domain ($A' = Q^{-1} \circ A \circ Q$). More generally, for $L : V \rightarrow W$ we may use *different* bases on the domain and co-domain. In fact we have essentially already done this in our rank-nullity theorem.

Namely suppose we have a linear map $L : V \rightarrow W$ from an n -dimensional vector space to an m -dimensional vector space with $\text{rank}(L) = r$. Choose some basis $\mathbf{v}_1, \dots, \mathbf{v}_{n-r}$ for $\ker(L)$ and extend to a basis $\mathbf{v}'_1, \dots, \mathbf{v}'_r, \mathbf{v}_1, \dots, \mathbf{v}_{n-r}$ of V . Then (exercise: prove it!) the vectors $\mathbf{w}'_1 = L(\mathbf{v}'_1), \dots, \mathbf{w}'_r = L(\mathbf{v}'_r) \in W$ are independent. Thus they can be extended to a basis $\mathbf{w}'_1, \dots, \mathbf{w}'_r, \mathbf{w}_1, \dots, \mathbf{w}_{m-r}$ of W .

Then, in these bases of V, W ‘adapted’ to L :

$$\mathbf{v}'_1, \dots, \mathbf{v}'_r, \mathbf{v}_1, \dots, \mathbf{v}_{n-r}$$

$$\mathbf{w}'_1, \dots, \mathbf{w}'_r, \mathbf{w}_1, \dots, \mathbf{w}_{m-r}$$

we have the simple matrix representation of L from $\mathbb{R}^n \rightarrow \mathbb{R}^m$ via matrix multiplication by the $m \times n$

$$\begin{pmatrix} I_{r \times r} & 0_{r \times n'} \\ 0_{m' \times r} & 0_{m' \times n'} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

setting $n' = n - r$, $m' = m - r$, $I_{r \times r}$ for the $r \times r$ identity matrix, and $0_{k \times l}$ for a $k \times l$ matrix all of whose entries are zeroes (exercise: check!).