Linear Algebra and Geometry

Informal notes on change of basis in a vector space

1 Matrices associated to a linear map

Let V and W be two vector spaces over a field $K \in \{\mathbb{R}, \mathbb{C}\}$, both finitely generated, $\dim_K(V) = n$, $\dim_K(W) = m$, having bases $\mathcal{B} = (v_1, \ldots, v_n)$ and $\mathcal{D} = (w_1, \ldots, w_m)$ respectively. Let

$$f : V \to W$$

be a linear map. **The matrix associated to f with respect to the bases $\mathcal{B}$ and $\mathcal{D}$, denoted by $M^{\mathcal{B}}_{\mathcal{D}}(f)$, is the $m \times n$ matrix whose j-th column is formed by the coordinates of the vector $f(v_j) \in W$ with respect to the basis $\mathcal{D}$, i.e. $[f(v_j)]_{\mathcal{D}}$.**

Explicitly, if $f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \ldots + a_{mj}w_m$, then:

$$M^{\mathcal{B}}_{\mathcal{D}}(f) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

Once we have defined the matrix associated to f, given any $v = x_1v_1 + \ldots + x_nv_n \in V$, if $f(v) = y_1w_1 + \ldots + y_mw_m \in W$, then:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = M^{\mathcal{B}}_{\mathcal{D}}(f) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

that is:

$$[f(v)]_{\mathcal{D}} = M^{\mathcal{B}}_{\mathcal{D}}(f)[v]_{\mathcal{B}}.$$

If U is another vector space over K, finitely generated of $\dim_K(U) = s$ and with a basis $\mathcal{U} = (u_1, \ldots, u_s)$, and

$$g : U \to V$$

is a linear map that can be composed with f, so that the linear map

$$\overbrace{U \xrightarrow[g]{} V \xrightarrow[f]{} W}^{f \circ g},$$

is well defined, then:

$$M^{\mathcal{U}}_{\mathcal{D}}(f \circ g) = M^{\mathcal{B}}_{\mathcal{D}}(f) M^{\mathcal{U}}_{\mathcal{B}}(g).$$

As a consequence, if $f : V \to V$ is an invertible endomorphism, and $\mathcal{B}$ and $\mathcal{D}$ are two bases of V, then:

$$M^{\mathcal{D}}_{\mathcal{B}}(f^{-1}) = (M^{\mathcal{B}}_{\mathcal{D}}(f))^{-1}$$

Moreover, if the map f coincides with the multiplication on the left by a matrix A, meaning if f is of the form μ_A, then $M^{\mathcal{B}}_{\mathcal{D}}(\mu_A) = A$, as expected.

Example 1. Let $\mathcal{B} = (v_1, v_2, v_3, v_4)$ and $\mathcal{D} = (w_1, w_2, w_3)$ be bases of $\mathbb{R}^4$ and $\mathbb{R}^3$ respectively, and let $f : \mathbb{R}^4 \to \mathbb{R}^3$ be the linear map defined by the equations:

$$\begin{aligned} f(v_1) &= 2w_1 - w_2 + 5w_3 \\ f(v_2) &= -w_2 - w_3 \\ f(v_3) &= w_1 + 3w_2 + 2w_3 \\ f(v_4) &= -w_1 + w_2 \end{aligned}$$

Let's compute the associated matrix $M^{\mathcal{B}}_{\mathcal{D}}(f)$.

By definition, $M^{\mathcal{B}}_{\mathcal{D}}(f)$ is the 3×4 matrix whose columns are the coordinates of the vectors $f(v_j)$ with respect to the w_is, and we already have this information. Indeed $f(v_1) = 2w_1 - w_2 + 5w_3$ means that the coordinates of $f(v_1)$ with respect to the basis $\mathcal{D}$ are $[f(v_1)]_{\mathcal{D}} = (2, -1, 5)$, and similarly for the other 3 elements of $\mathcal{B}$. Hence:

$$M^{\mathcal{B}}_{\mathcal{D}}(f) = \begin{pmatrix} 2 & 0 & 1 & -1 \\ -1 & -1 & 3 & 1 \\ 5 & -1 & 2 & 0 \end{pmatrix}. \qquad \checkmark$$

2 Change of coordinates matrix (or change of basis matrix)

A particularly important situation is when we deal with two different bases $\mathcal{B} = (v_1, \dots, v_n)$ and $\mathcal{D} = (w_1, \dots, w_n)$ of the same vector space V, and $f : V \to V$ is the identity map $f = \mathrm{id}_V$. In this case **the matrix $M^{\mathcal{B}}_{\mathcal{D}}(id_V)$ is called matrix of basis change (or change of basis) from the basis $\mathcal{B}$ to the basis $\mathcal{D}$.** By definition, $M^{\mathcal{B}}_{\mathcal{D}}(id_V)$ **is the matrix whose j-th column is formed by the coordinates of the vector $id_V(v_j) = v_j$ with respect to the basis $\mathcal{D}$, i.e. $[v_j]_{\mathcal{D}}$.**

Example 2. Let's look at the same example as before, and suppose that we know that the coordinates of the elements of the basis $\mathcal{D}$ of $\mathbb{R}^3$ with respect to the standard basis $\mathcal{C}$ are:

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \quad w_1 = \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}.$$

Let us write down the two matrices of basis change: $M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})$, taking us from the basis $\mathcal{D}$ to the canonical basis $\mathcal{C}$, and $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$ that takes us from $\mathcal{C}$ to $\mathcal{D}$. The first one is basically already there; indeed $M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})$ is the matrix whose columns are the coordinates of the w_js with respect to the canonical basis, that is:

$$M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}) = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 5 \end{pmatrix}.$$

To compute $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$ we can go in two different ways, that are equivalent: the first method is to express the vectors e_is as combinations of the w_js, meaning that we find explicitly the coordinates $[e_j]_{\mathcal{D}}$, that will form the columns of $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$. Otherwise we can remember that $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3}) = (M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}))^{-1}$, and then take the inverse of the matrix above:

$$M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3}) = (M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}))^{-1} = \frac{1}{3}\begin{pmatrix} -2 & 5 & -1 \\ -5 & 5 & -1 \\ 2 & -2 & 1 \end{pmatrix}. \qquad \checkmark$$

3 Let's put all together

Going back to the general case of two vector spaces V and W having bases $\mathcal{B}$ and $\mathcal{D}$ respectively, let us suppose that we have a linear map

$$f : (V, \mathcal{B}) \to (W, \mathcal{D}),$$

where this notation means that f sends an element of the space V having a fixed basis $\mathcal{B}$, into an element of the space W with a fixed basis $\mathcal{D}$. What happens to the matrix associated to f if we change basis in V? and if we change basis in W?

To express the matrix associated to f with respect to the new bases, say $\mathcal{E}$ for V and $\mathcal{F}$ for W, we proceed by composing linear mappings in the following way:

$$M_{\mathcal{F}}^{\mathcal{E}}(f) : (V, \mathcal{E}) \xrightarrow{M_{\mathcal{B}}^{\mathcal{E}}(id_V)} (V, \mathcal{B}) \xrightarrow{M_{\mathcal{D}}^{\mathcal{B}}(f)} (W, \mathcal{D}) \xrightarrow{M_{\mathcal{F}}^{\mathcal{D}}(id_W)} (W, \mathcal{F})$$

or, equivalently, in the diagram:

$$\begin{array}{ccc}
(V,\mathcal{B}) & \xrightarrow{M^{\mathcal{B}}_{\mathcal{D}}(f)} & (W,\mathcal{D}) \\
{\scriptstyle M^{\mathcal{E}}_{\mathcal{B}}(id_V)}\uparrow & & \downarrow{\scriptstyle M^{\mathcal{D}}_{\mathcal{F}}(id_W)} \\
(V,\mathcal{E}) & \xrightarrow[M^{\mathcal{E}}_{\mathcal{F}}(f)]{} & (W,\mathcal{F})
\end{array}$$

Now we only need to remember that the composition of linear mappings corresponds to the product of the associated matrices, and we are done:

$$M^{\mathcal{E}}_{\mathcal{F}}(f) = M^{\mathcal{D}}_{\mathcal{F}}(id_W)\ M^{\mathcal{B}}_{\mathcal{D}}(f)\ M^{\mathcal{E}}_{\mathcal{B}}(id_V).$$

Example 3. In *Example 1* we computed the matrix $M^{\mathcal{B}}_{\mathcal{D}}(f)$ associated to the linear map

$$f : (\mathbb{R}^4, \mathcal{B}) \to (\mathbb{R}^3, \mathcal{D}).$$

In *Example 2* instead, we computed the two matrices of change of basis in $\mathbb{R}^3$ $M^{\mathcal{D}}_{\mathcal{C}}(id_{\mathbb{R}^3})$ and $M^{\mathcal{C}}_{\mathcal{D}}(id_{\mathbb{R}^3})$.

Now let's compute the matrix $M^{\mathcal{B}}_{\mathcal{C}}(\tilde{f})$ associated to the linear map

$$\tilde{f} : (\mathbb{R}^4, \mathcal{B}) \to (\mathbb{R}^3, \mathcal{C}).$$

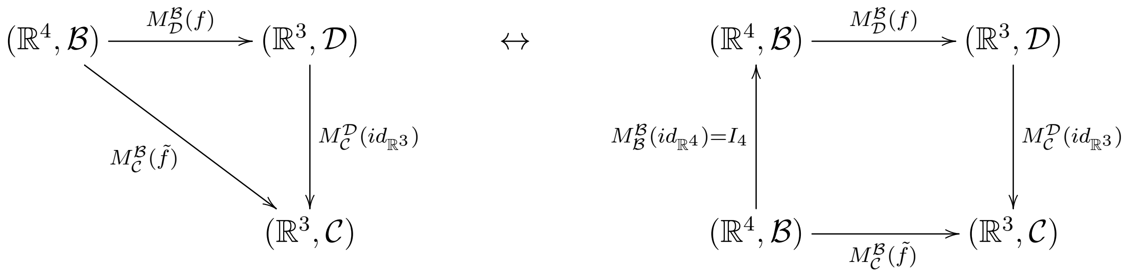
In our particular situation, the diagram above becomes a triangle instead of a square, because we only change basis in the codomain:

$$\begin{array}{ccc}
(\mathbb{R}^4,\mathcal{B}) & \xrightarrow{M^{\mathcal{B}}_{\mathcal{D}}(f)} & (\mathbb{R}^3,\mathcal{D}) \\
 & {\scriptstyle M^{\mathcal{B}}_{\mathcal{C}}(\tilde{f})}\searrow & \downarrow{\scriptstyle M^{\mathcal{D}}_{\mathcal{C}}(id_{\mathbb{R}^3})} \\
 & & (\mathbb{R}^3,\mathcal{C}).
\end{array}$$

Again, the composition of maps corresponds to the product of matrices, and hence:

$$M^{\mathcal{B}}_{\mathcal{C}}(\tilde{f}) = M^{\mathcal{D}}_{\mathcal{C}}(id_{\mathbb{R}^3})\ M^{\mathcal{B}}_{\mathcal{D}}(f) = \begin{pmatrix} 3 & 1 & -2 & -2 \\ 7 & -1 & 3 & -1 \\ 23 & -7 & 16 & 2 \end{pmatrix}. \qquad \checkmark$$

If going from the square diagram to the triangular one confuses you, you can always look at the latter as a square where one of the sides is given by the identity matrix, since on $\mathbb{R}^4$ we are not performing any change of basis:



Obviously $M_{\mathcal{C}}^{\mathcal{B}}(\tilde{f}) = M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})\ M_{\mathcal{D}}^{\mathcal{B}}(f) = M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})\ M_{\mathcal{D}}^{\mathcal{B}}(f)\, I_4$.

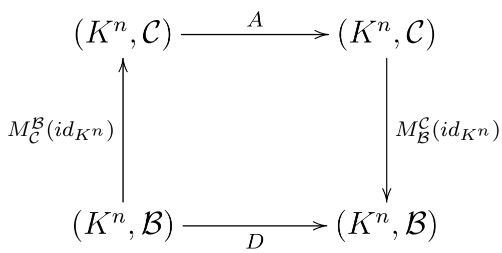
4 A special case: diagonalization

Diagonalizing a matrix (when possible) is exactly the same process that we described above. We start from a matrix $A \in K^{n,n}$, or, if you prefer, from the linear mapping (endomorphism)

$$f = \mu_A : K^n \to K^n,$$

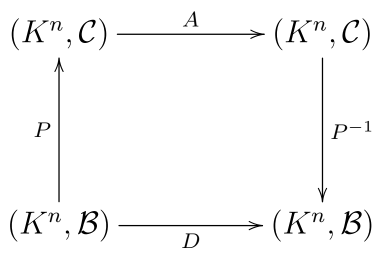
and hence, using the same notation as above: $V \cong W \cong K^n$, the bases $\mathcal{B}$ and $\mathcal{D}$ are both equal to the canonical basis $\mathcal{C}$, and finally the matrix $M_{\mathcal{C}}^{\mathcal{C}}(f) = A$. **We are looking for a new basis $\mathcal{B}$ such that $M_{\mathcal{B}}^{\mathcal{B}}(f)$ is a diagonal matrix D.**

Assume that the matrix A is diagonalizable, meaning that it is possible to find a basis $\mathcal{B} = (v_1, \ldots, v_n)$ formed by eigenvectors $v_1, \ldots, v_n$ relative to eigenvalues $\lambda_1, \ldots, \lambda_n$ (not necessarily distinct) respectively. If we apply the change of basis, the diagram we had above can be written as:



By definition, the matrix $M_{\mathcal{C}}^{\mathcal{B}}(id_{K^n})$ is the matrix whose columns are the eigenvectors v_js, whereas the matrix $M_{\mathcal{B}}^{\mathcal{C}}(id_{K^n})$ is its inverse: $M_{\mathcal{B}}^{\mathcal{C}}(id_{K^n}) = (M_{\mathcal{C}}^{\mathcal{B}}(id_{K^n}))^{-1}$.

For the sake of simplicity, let us call $P = M_{\mathcal{C}}^{\mathcal{B}}(id_{K^n})$, so that the diagram becomes:



and hence

$$D = P^{-1}AP.$$

We remark that

$$D = P^{-1}AP \quad \Leftrightarrow \quad PD = AP \quad \Leftrightarrow \quad PDP^{-1} = A.$$

Example 4. Let us diagonalize the matrix

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Its characteristic polynomial (expanding along the second row) is:

$$p_A(t) = \det(A - tI_3) = \det \begin{pmatrix} 1-t & 0 & 4 \\ 0 & 2-t & 0 \\ 1 & 0 & 1-t \end{pmatrix} = (2-t)[(1-t)^2 - 4],$$

having three distinct roots $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_3 = -1$: Therefore, the matrix A is diagonalizable for sure.

Let's find a basis of eigenvectors and perform the change of basis.

The three eigenspaces are:

$$E_A(2) = \mathrm{Ker}(A - 2I_3) = \mathrm{Ker} \begin{pmatrix} -1 & 0 & 4 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \mathcal{L}\big((0,1,0)\big),$$

$$E_A(3) = \mathrm{Ker}(A - 3I_3) = \mathrm{Ker} \begin{pmatrix} -2 & 0 & 4 \\ 0 & -1 & 0 \\ 1 & 0 & -2 \end{pmatrix} = \mathcal{L}\big((2,0,1)\big),$$

$$E_A(-1) = \mathrm{Ker}(A + I_3) = \mathrm{Ker} \begin{pmatrix} 2 & 0 & 4 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix} = \mathcal{L}\big((-2,0,1)\big).$$

To conclude, we verify that:

$$P^{-1}AP = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} \\ -\frac{1}{4} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D. \qquad \checkmark$$

5 An even more special case: symmetric matrices

If we want to diagonalize a symmetric matrix, the Spectral Theorem not only guarantees that we can always do it on the field of real numbers, but we can also find a diagonalizing matrix P which is orthogonal, i.e. such that $P^{-1} = {}^tP$. This simplifies even more our task, since we don't have to struggle to invert P.

Be careful that, even if eigenvectors (of a symmetric matrix) relative to distinct eigenvalues are automatically orthogonal, this is not true for eigenvectors relative to a same eigenvalue with geometric multiplicity > 1: in this latter case we need to find an orthogonal basis for the eigenspace. Usually one can "eyeball" the computation; if that's not the case, the Gram-Schmidt's algorithm comes to the rescue.

Example 5. Let's diagonalize the symmetric matrix

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

and find a diagonalizing orthonormal basis.

Its characteristic polynomial is:

$$p_A(t) = \det(A - tI_3) = \det \begin{pmatrix} 1-t & 2 & 2 \\ 2 & 1-t & 2 \\ 2 & 2 & 1-t \end{pmatrix} = \cdots = -(t-5)(t+1)^2,$$

having two distinct roots $\lambda_1 = 5$ and $\lambda_2 = -1$, where this last one has algebraic multiplicity 2 (but we are not worried about it, because the Spectral Theorem ensures that its geometric multiplicity will be 2).

Let us find an orthonormal basis made up of eigenvectors and perform the change of basis. The two eigenspaces are:

$$E_A(5) = \mathrm{Ker}(A - 5I_3) = \mathrm{Ker} \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} = \mathcal{L}\big((1,1,1)\big),$$

$$E_A(-1) = \mathrm{Ker}(A + I_3) = \mathrm{Ker} \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} = \mathcal{L}\big((-1,1,0),(-1,0,1)\big).$$

As expected, we have 3 independent eigenvectors forming a basis. But we don't have an orthonormal basis yet! The only property we have "for free" is that the eigenvector v_1 relative to the eigenvalue $\lambda_1 = 5$ is orthogonal to both eigenvectors v_2 and v_3 relative to the eigenvalue $\lambda_2 = -1$. Let us transform the basis $\mathcal{B} = (v_1, v_2, v_3)$ into an orthonormal basis $\mathcal{B}' = (w_1, w_2, w_3)$.

To get w_1 it is enough to normalize v_1, so

$$w_1 = \frac{v_1}{|v_1|} = \Big(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big).$$

To get the other two vector we apply Gram-Schmidt's algorithm to the basis (v_1, v_2) of $E_A(-1)$:

$$\begin{aligned} u_2 &= v_2 = (-1, 1, 0), \\ w_3 &= v_3 - \frac{v_3 \cdot u_2}{|u_2|^2} u_2 = v_3 - \frac{1}{2} u_2 = \Big(-\frac{1}{2}, -\frac{1}{2}, \ 1\Big), \end{aligned}$$

and from this we eventually deduce:

$$\begin{aligned} w_2 &= \frac{u_2}{|u_2|} = \Big(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \ 0\Big), \\ w_3 &= \frac{u_3}{|u_3|} = \Big(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\Big). \end{aligned}$$

To conclude, let's verify that $P^{-1}AP = {}^tPAP = D$: indeed, one has

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \qquad \checkmark$$

N.B. If something is not clear/you don't understand it, or if you find a typo, please let me know!