Linear Transformations

Fundamentals of Robotics, Lecture 2

by Sergei Savin

Fall 2022

The goal of this lecture is to help you perform change of basis correctly and easily. To do it we will try to think in terms of *abstract vectors* rather than vector *coordinates*.

The motivation is simple. If you have to perform a transformation while thinking about a vector in terms of its coordinates, you have to keep in mind three entities: vector, the basis in which it is expressed, giving you the coordinates, and the transformation. Giving up on coordinates simplifies it to only two entities.

First, a small exercise in abstract thinking and linear algebra.

Consider a vector $\mathbf{a} \in \mathscr{L}$, where $\mathscr{L}$ is a linear space.

We are very used to think about the vector space $\mathbb{R}^n$ where vectors are essentially columns of numbers. But you are also familiar with vectors as arrows. For today's lecture, let us think about vectors as arrows, or purely abstract objects.

Consider vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n \in \mathscr{L}$. If $\mathbf{a} \in \mathrm{span}(\mathbf{v}_1,\ ....,\ \mathbf{v}_n)$, it means there are coordinates α_1, α_2, ..., $\alpha_n \in \mathbb{R}$ that allow us to express $\mathbf{a}$ in terms of vectors $\mathbf{v}_i$:

$$\mathbf{a} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + ... + \alpha_n \mathbf{v}_n \tag{1}$$

We can define linear operations that map vectors from $\mathscr{L}$ to itself. For example, let $\mathcal{M}$ be such operation: $\mathcal{M} : \mathscr{L} \to \mathscr{L}$.

Remember that $\mathcal{M}$ is linear:

$$\mathcal{M}(\beta_1 \mathbf{w}_1 + \beta_2 \mathbf{w}_2) = \beta_1 \mathcal{M}(\mathbf{w}_1) + \beta_2 \mathcal{M}(\mathbf{w}_2) \tag{2}$$

where $\mathbf{w}_1, \mathbf{w}_2 \in \mathscr{L}$ and $\beta_1, \beta_2 \in \mathbb{R}$.

This is useful, because for our vector $\mathbf{a} = \alpha_1 \mathbf{v}_1 + ... + \alpha_n \mathbf{v}_n$ the operator $\mathcal{M}$ then works as follows:

$$\mathcal{M}(\mathbf{a}) = \alpha_1 \mathcal{M}(\mathbf{v}_1) + ... + \alpha_n \mathcal{M}(\mathbf{v}_n) \tag{3}$$

Why is the fact $\mathcal{M}(\mathbf{a}) = \alpha_1 \mathcal{M}(\mathbf{v}_1) + ... + \alpha_n \mathcal{M}(\mathbf{v}_n)$ useful? Because instead of thinking about what $\mathcal{M}$ does to any possible $\mathbf{a} \in \mathscr{L}$, we think about what it does to n particular vectors - $\mathbf{v}_1$, ... $\mathbf{v}_n$.

Let us define $\mathbf{u}_1 = \mathcal{M}(\mathbf{v}_1)$, ..., $\mathbf{u}_n = \mathcal{M}(\mathbf{v}_n)$.

If $\mathbf{v}_1$, ... $\mathbf{v}_n$ form a basis in $\mathscr{L}$, then all vectors $\mathbf{u}_i$ can also be represented as linear combinations of $\mathbf{v}_1$, ... $\mathbf{v}_n$. For example:

$$\begin{aligned}
\mathbf{u}_1 &= m_{11}\mathbf{v}_1 + m_{21}\mathbf{v}_2 + ... + m_{n1}\mathbf{v}_n \\
\mathbf{u}_2 &= m_{12}\mathbf{v}_1 + m_{22}\mathbf{v}_2 + ... + m_{2n}\mathbf{v}_n \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad ... \\
\mathbf{u}_n &= m_{1n}\mathbf{v}_1 + m_{2n}\mathbf{v}_2 + ... + m_{nn}\mathbf{v}_n
\end{aligned}$$

Let us define $\mathbf{b} = \mathcal{M}(\mathbf{a})$; since $\mathbf{v}_1$, ... $\mathbf{v}_n$ form a basis in $\mathscr{L}$, we can represent $\mathbf{b}$ as:

$$\mathbf{b} = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + ... + \beta_n \mathbf{v}_n \tag{4}$$

How can we find coordinates β_1, β_2, ..., β_n if we know α_1, α_2, ..., α_n?

$$\mathcal{M}(\mathbf{a}) = \alpha_1 \mathcal{M}(\mathbf{v}_1) + ... + \alpha_n \mathcal{M}(\mathbf{v}_n) =$$
$$\alpha_1(m_{11}\mathbf{v}_1 + m_{21}\mathbf{v}_2 + ... + m_{n1}\mathbf{v}_n) + ...+$$
$$\alpha_n(m_{1n}\mathbf{v}_1 + m_{2n}\mathbf{v}_2 + ... + m_{nn}\mathbf{v}_n) =$$
$$(\alpha_1 m_{11} + ... + \alpha_n m_{1n})\mathbf{v}_1 + ... + (\alpha_1 m_{n1} + ... + \alpha_n m_{nn})\mathbf{v}_n =$$
$$\beta_1 \mathbf{v}_1 + ... + \beta_n \mathbf{v}_n$$

which is to say:

$$\beta_1 = \alpha_1 m_{11} + ... + \alpha_n m_{1n}$$
$$\beta_2 = \alpha_1 m_{21} + ... + \alpha_n m_{2n}$$
$$...$$
$$\beta_n = \alpha_1 m_{n1} + ... + \alpha_n m_{nn}$$

But this:

$$\beta_1 = \alpha_1 m_{11} + ... + \alpha_n m_{1n}$$
$$\beta_2 = \alpha_1 m_{21} + ... + \alpha_n m_{2n}$$
$$...$$
$$\beta_n = \alpha_1 m_{n1} + ... + \alpha_n m_{nn}$$

is the same as matrix multiplication:

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ ... \\ \beta_n \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & ... & m_{1n} \\ m_{21} & m_{22} & ... & m_{2n} \\ ... & ... & ... & ... \\ m_{n1} & m_{n2} & ... & m_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ ... \\ \alpha_n \end{bmatrix} \tag{5}$$

This matrix completely describes what the operator $\mathcal{M}$ does, using a specific basis.

What did we learn?

- Vectors and operators exist independent of coordinate representation.
- Coordinates are associates with bases.
- Coordinates describe both vectors and operators.

Now assume that aside from basis $\mathbf{v}_1$, ... $\mathbf{v}_n$, we also have a basis $\mathbf{w}_1$, ... $\mathbf{w}_n$. If we know coordinates of $\mathbf{a}$ in $\mathbf{v}_1$, ... $\mathbf{v}_n$, can we find its coordinates in $\mathbf{w}_1$, ... $\mathbf{w}_n$?

$$\mathbf{a} = \gamma_1 \mathbf{w}_1 + \gamma_2 \mathbf{w}_2 + ... + \gamma_n \mathbf{w}_n \tag{6}$$

First, we need coordinates of vectors $\mathbf{w}_i$ in our old basis:

$$\begin{aligned}
\mathbf{w}_1 &= c_{11}\mathbf{v}_1 + c_{21}\mathbf{v}_2 + ... + c_{n1}\mathbf{v}_n \\
\mathbf{w}_2 &= c_{12}\mathbf{v}_1 + c_{22}\mathbf{v}_2 + ... + c_{2n}\mathbf{v}_n \\
&\qquad\qquad\qquad\qquad\qquad\qquad ... \\
\mathbf{w}_n &= c_{1n}\mathbf{v}_1 + c_{2n}\mathbf{v}_2 + ... + c_{nn}\mathbf{v}_n
\end{aligned}$$

Then, from $\mathbf{a} = \gamma_1 \mathbf{w}_1 + \gamma_2 \mathbf{w}_2 + ... + \gamma_n \mathbf{w}_n$ we get:

$$\begin{aligned} \mathbf{a} = \gamma_1(c_{11}\mathbf{v}_1 + c_{21}\mathbf{v}_2 + ... + c_{n1}\mathbf{v}_n) + ...+ \\ \gamma_n(c_{1n}\mathbf{v}_1 + c_{2n}\mathbf{v}_2 + ... + c_{nn}\mathbf{v}_n) = \\ (\gamma_1 c_{11} + ... + \gamma_n c_{1n})\mathbf{v}_1 + ... + (\gamma_1 c_{n1} + ... + \gamma_n c_{nn})\mathbf{v}_n = \\ \alpha_1 \mathbf{v}_1 + ... + \alpha_n \mathbf{v}_n \end{aligned}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ ... \\ \alpha_n \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & ... & c_{1n} \\ c_{21} & c_{22} & ... & c_{2n} \\ ... & ... & ... & ... \\ c_{n1} & c_{n2} & ... & c_{nn} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ ... \\ \gamma_n \end{bmatrix} \tag{7}$$

We see that to find γ_i we will need to invert the matrix, unlike in the previous example. However, that is only because vectors $\mathbf{w}_i$ were expressed in the basis $\mathbf{v}_1$, ... $\mathbf{v}_n$, and not the other way around.

Things to note here:

- In Robotics we sometimes talk about "active" rotations (that rotate a vector, result is expressed in the same basis) and "passive" rotations (is just a change of basis). It can become intensely confusing if "active" operations are performed but the result is expressed in a different basis.
- Our approach so far does not require this terminology and abstractions; it works with the most basic principles of linear algebra. Thus, it is useful for checking your results.

Now we will consider a specific linear space - $\mathbb{R}^n$. This will noticeably increase the risk of confusion:

- We cannot tell the difference between a vector in $\mathbb{R}^n$ and coordinates of a vector in $\mathbb{R}^n$.
- A basis in $\mathbb{R}^n$ is just a n-by-n matrix; same is true about an operator.

Our aim now is to resist the temptation to treat coordinates and vectors as indistinct.

Assume that columns of a matrix $\mathbf{V} = [\mathbf{v}_1, \ ... \ \mathbf{v}_n]$ form an orthonormal basis $\mathcal{V}$ in $\mathbb{R}^n$. Orthonormal means - all columns of $\mathbf{V}$ have norm 1, and they are all mutually orthogonal.

Consider a vector $\mathbf{a} = \alpha_1 \mathbf{v}_1 + ... + \alpha_n \mathbf{v}_n$. We can find coordinates α_i as follows:

$$\alpha_i = \mathbf{v}_i \cdot \mathbf{a} \tag{8}$$

which is true because $\mathbf{v}_i \cdot \mathbf{v}_i = 1$ and for all $i \neq j$ $\mathbf{v}_i \cdot \mathbf{v}_j = 0$. If the basis was not orthonormal, but we could find dot products, we could still compute coordinates - figure out how!

Knowing that $\alpha_i = \mathbf{v}_i \cdot \mathbf{a}$ we can find coordinates in one matrix multiplication:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}^\top \mathbf{a} = \mathbf{V}^\top \mathbf{a} \tag{9}$$

In general, if $\mathbf{r} \in \mathbb{R}^n$ is a vector, $\mathbf{B}$ is a basis and $\mathbf{p} \in \mathbb{R}^n$ are coordinates of $\mathbf{r}$ in $\mathbf{B}$, then the following relations hold:

$$\mathbf{r} = \mathbf{B}\mathbf{p} \tag{10}$$

$$\mathbf{p} = \mathbf{B}^\top \mathbf{r} \tag{11}$$

Any matrix $\mathbf{B}$ whose columns form an orthonormal basis in $\mathbb{R}^n$ is called an orthonormal matrix. Such matrices have properties:

$$\mathbf{B}^{-1} = \mathbf{B}^\top \tag{12}$$

$$\mathbf{B}^\top \mathbf{B} = \mathbf{I} \tag{13}$$

These follow directly from the definition of orthonormal matrices. It can also be used to directly prove the results on the previous slide.

Let $\mathbf{r}$ be expressed in the orthonormal basis $\mathcal{A}$ formed by columns of $\mathbf{A}$, with coordinate vector ${}^{\mathcal{A}}\mathbf{r}$. Let us find coordinates of $\mathbf{r}$ in the orthonormal basis $\mathcal{B}$ formed by columns of $\mathbf{B}$.

As long as we know $\mathbf{r}$, the solution is trivial: ${}^{\mathcal{B}}\mathbf{r} = \mathbf{B}^\top \mathbf{r}$. If $\mathbf{r}$ is not known, then we can first find it:

$$\mathbf{r} = \mathbf{A}\ {}^{\mathcal{A}}\mathbf{r} \tag{14}$$

$${}^{\mathcal{B}}\mathbf{r} = \mathbf{B}^\top \mathbf{A}\ {}^{\mathcal{A}}\mathbf{r} \tag{15}$$

We could memorize that in this scenario the relation between coordinates is given by matrix multiplication $\mathbf{B}^\top \mathbf{A}$, but it is better to just understand the process.

Let us look at the expression below one more time:

$$\mathbf{r} = \mathbf{A}\ {}^{\mathcal{A}}\mathbf{r} \tag{16}$$

$${}^{\mathcal{B}}\mathbf{r} = \mathbf{B}^{\top}\mathbf{A}\ {}^{\mathcal{A}}\mathbf{r} \tag{17}$$

In writing kinematics equations, you will feel great uneasiness about vectors like **r**, which have no indicators as to in which basis they are expressed. But remember - **r** is not a stack of coordinates, it is an element of the vector space $\mathscr{L}$. The vector space just happens to be $\mathbb{R}^n$, which leads us to desire to think only of coordinates, not abstract vectors. If you insist of thinking about **r** as coordinates - its basis is identity matrix. That is also the basis in which columns of both **A** and **B** are expressed.

Let $\mathbf{r}$ be expressed in the orthonormal basis $\mathcal{A}$ formed by columns of $\mathbf{A}$. We know that columns of $\mathbf{A}$ expressed in the orthonormal basis $\mathcal{C}$ are given as ${}^{\mathcal{C}}\mathbf{A}$. The basis $\mathcal{C}$ is formed by the columns of matrix $\mathbf{C}$. We know that columns of $\mathbf{C}$ expressed in the orthonormal basis $\mathcal{W}$ are given as ${}^{\mathcal{W}}\mathbf{C}$.

Given basis $\mathcal{B}$ formed by the columns of matrix $\mathbf{B}$, expressed in the orthonormal basis $\mathcal{W}$ as ${}^{\mathcal{W}}\mathbf{B}$, find transformation $\mathbf{T}$ that maps coordinates of $\mathbf{r}$ in $\mathcal{A}$ to coordinates of $\mathbf{r}$ in $\mathcal{B}$.

Sounds like a nightmare, doesn't it?

First, we find coordinates of $\mathbf{r}$ in $\mathcal{A}$:

$$ {}^{\mathcal{A}}\mathbf{r} = \mathbf{A}^{\top}\mathbf{r} \tag{18} $$

But we do not know $\mathbf{A}$, we only know ${}^{\mathcal{C}}\mathbf{A}$. Well, this knowing the coordinates of the columns, finding the columns themselves is easy:

$$ \mathbf{A} = \mathbf{C} \ {}^{\mathcal{C}}\mathbf{A} \tag{19} $$

Same is true for $\mathbf{C}$ - we do not know it, but we have the coordinates:

$$ \mathbf{C} = \mathbf{W} \ {}^{\mathcal{W}}\mathbf{C} \tag{20} $$

$$ \mathbf{A} = \mathbf{W} \ {}^{\mathcal{W}}\mathbf{C} \ {}^{\mathcal{C}}\mathbf{A} \tag{21} $$

$$ {}^{\mathcal{A}}\mathbf{r} = (\mathbf{W} \ {}^{\mathcal{W}}\mathbf{C} \ {}^{\mathcal{C}}\mathbf{A})^{\top}\mathbf{r} \tag{22} $$

$$ {}^{\mathcal{A}}\mathbf{r} =^{\mathcal{C}} \mathbf{A}^{\top} \ {}^{\mathcal{W}}\mathbf{C}^{\top}\mathbf{W}^{\top}\mathbf{r} \tag{23} $$

Next, we find coordinates of $\mathbf{r}$ in $\mathcal{B}$:

$$^{\mathcal{B}}\mathbf{r} = \mathbf{B}^\top \mathbf{r} \tag{24}$$

As before, we do not know $\mathbf{B}$, we only know $^{\mathcal{W}}\mathbf{B}$:

$$\mathbf{B} = \mathbf{W} \; ^{\mathcal{W}}\mathbf{B} \tag{25}$$

$$^{\mathcal{B}}\mathbf{r} = (\mathbf{W} \; ^{\mathcal{W}}\mathbf{B})^\top \mathbf{r} \tag{26}$$

$$^{\mathcal{B}}\mathbf{r} = ^{\mathcal{W}} \mathbf{B}^\top \mathbf{W}^\top \mathbf{r} \tag{27}$$

Now, what is the relation between $^{\mathcal{A}}\mathbf{r} = ^{\mathcal{C}} \mathbf{A}^\top \; ^{\mathcal{W}}\mathbf{C}^\top \mathbf{W}^\top \mathbf{r}$ and $^{\mathcal{B}}\mathbf{r} = ^{\mathcal{W}} \mathbf{B}^\top \mathbf{W}^\top \mathbf{r}$?

$$^{\mathcal{W}}\mathbf{C} \; ^{\mathcal{C}}\mathbf{A} \; ^{\mathcal{A}}\mathbf{r} = \mathbf{W}^\top \mathbf{r} \tag{28}$$

$$^{\mathcal{W}}\mathbf{B} \; ^{\mathcal{B}}\mathbf{r} = \mathbf{W}^\top \mathbf{r} \tag{29}$$

$$^{\mathcal{W}}\mathbf{C} \; ^{\mathcal{C}}\mathbf{A} \; ^{\mathcal{A}}\mathbf{r} = ^{\mathcal{W}} \mathbf{B} \; ^{\mathcal{B}}\mathbf{r} \tag{30}$$

From the last expression:

$$ {}^{\mathcal{W}}\mathbf{C} \ \ {}^{\mathcal{C}}\mathbf{A} \ \ {}^{\mathcal{A}}\mathbf{r} = {}^{\mathcal{W}}\mathbf{B} \ \ {}^{\mathcal{B}}\mathbf{r} \tag{31} $$

...we get the solution:

$$ {}^{\mathcal{B}}\mathbf{r} = {}^{\mathcal{W}}\mathbf{B}^{\top} \ {}^{\mathcal{W}}\mathbf{C} \ \ {}^{\mathcal{C}}\mathbf{A} \ \ {}^{\mathcal{A}}\mathbf{r} \tag{32} $$

$$ \mathbf{T} = {}^{\mathcal{W}}\mathbf{B}^{\top} \ {}^{\mathcal{W}}\mathbf{C} \ \ {}^{\mathcal{C}}\mathbf{A} \tag{33} $$

Notice that the process is simple, requiring no geometric intuition.

Final advice

Remember that applied Robotics is not a sprint: there are no bonus point for solving a change of basis problem faster. Instead, the goal is to be precise, to verify your results.

Do not shy away from methods that require you to fill a page with intermediate derivations if they are clear and lead to correct answers.

Thank you!

Lecture slides are available via Moodle.

You can help improve these slides at:
github.com/SergeiSa/Fundumentals-of-robotics-2022

Check Moodle for additional links, videos, textbook suggestions.