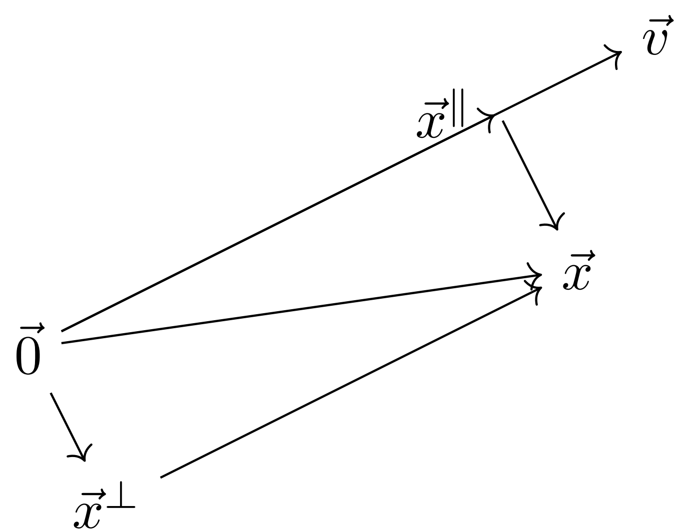
Orthogonal projection

Earlier, we learned about orthogonal projection onto the line spanned by a vector $\vec{v}$:

Given a nonzero vector $\vec{v}$ in $\mathbb{R}^n$, we can decompose any vector $\vec{x}$ as $\vec{x}^{\parallel} + \vec{x}^{\perp}$, where $\vec{x}^{\parallel}$ is parallel to $\vec{v}$ and $\vec{x}^{\perp}$ is perpendicular.



We have $x^{\parallel} = \dfrac{\vec{v}\cdot\vec{x}}{\vec{v}\cdot\vec{v}}\vec{v}$. If $\vec{v}$ has length 1, $\vec{x}^{\parallel} = (\vec{v}\cdot\vec{x})\vec{v}$.

Now we want to project, not onto a line, but onto a subspace L. Here is the analogous claim:

Given a subspace L in $\mathbb{R}^n$, we can decompose any vector $\vec{x}$ as $\vec{x}^{\|} + \vec{x}^{\perp}$, where $\vec{x}^{\|}$ is in L and $x^{\perp}$ is in $L^{\perp}$.

Today, we'll talk about how to compute $x^{\|}$ and $x^{\perp}$, assuming that we have an orthonormal basis $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_k$ for L. In the next lecture, we'll talk about how to find such an orthonormal basis.

Next week, we will also mention a formula for $\vec{x}^{\|}$ directly, without going through the intermediate step of finding an orthonormal basis for L.

Let $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_k$ be an orthonormal basis for L.

Since $\vec{x}^{\|}$ is in L, we can write

$$\vec{x}^{\|} = c_1\vec{v}_1 + \cdots + c_k\vec{v_k}$$

for some scalars c_j. We then have

$$\vec{v}_j \cdot \vec{x} = \vec{v}_j \cdot \left(\vec{x}^{\|} + \vec{x}^{\perp}\right) = \vec{v}_j \cdot \vec{x}^{\|} + \vec{v}_j \cdot \vec{x}^{\perp} = \vec{v}_j \cdot \vec{x}^{\|} + 0 = \vec{v}_j \cdot \vec{x}^{\|}.$$

But we have

$$\vec{v}_j \cdot \vec{x}^{\|} = \vec{v}_j \cdot (c_1\vec{v}_1 + \cdots + c_k\vec{v}_k) = c_1 0 + c_2 0 + \cdots + c_j 1 + \cdots + c_k 0 = c_j.$$

Putting it all together,

$$c_j = \vec{x} \cdot \vec{v}_j.$$

So:

$$\vec{x}^{\|} = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + \cdots + (\vec{v}_k \cdot \vec{x})\vec{v}_k.$$

$$\vec{x}^{\parallel} = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + \cdots + (\vec{v}_k \cdot \vec{x})\vec{v}_k.$$

This is the generalization of $\vec{x}^{\parallel} = (\vec{v} \cdot \vec{x})\vec{v}$, when $\vec{v}$ is a unit vector and we are projecting onto the line $\mathrm{Span}(\vec{v})$.

Of course,

$$\vec{x}^{\perp} = \vec{x} - \vec{x}^{\parallel} = \vec{x} - (\vec{v}_1 \cdot \vec{x})\vec{v}_1 - \cdots - (\vec{v}_k \cdot \vec{x})\vec{v}_k.$$

$$\vec{x}^{\parallel} = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + \cdots + (\vec{v}_k \cdot \vec{x})\vec{v}_k.$$

There is a cute way to write this with matrices. Let

$$Q = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \\ | & | & \cdots & | \end{bmatrix}.$$

So $Q^T Q = \mathrm{Id}_k$.

Then

$$x^{\parallel} = QQ^T \vec{x}.$$

Let's see why:

$$QQ^T\vec{x} = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} - & \vec{v}_1 & - \\ - & \vec{v}_2 & - \\ & \vdots & \\ - & \vec{v}_k & - \end{bmatrix} \vec{x} = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \vec{v_1} \cdot \vec{x} \\ \vec{v_2} \cdot \vec{x} \\ \cdots \\ \vec{v_k} \cdot \vec{x} \end{bmatrix}$$

$$= (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + \cdots + (\vec{v}_k \cdot \vec{x})\vec{v}_k.$$

So, if Q is an $n \times k$ matrix with $Q^T Q = \mathrm{Id}_k$, then the columns of k are an orthonormal basis for some k-dimensional subspace L, and QQ^T is the orthogonal projection onto L.

At this point, I hope you are convinced that orthonormal bases are great. In the next lecture, we'll talk about how to find them.