Appendix: only orthogonal matrices preserve the scalar product

I'm including this proof for completeness, for those who are interested. I think the result is sufficiently plausible that we didn't need to prove it rigorously in the lecture.

Let

$$\mathbf{x}' = \mathsf{R}\mathbf{x}$$
$$\mathbf{y}' = \mathsf{R}\mathbf{y}.$$

We saw that *if* $\mathsf{R}^\mathsf{T}\mathsf{R} = \mathtt{I}$, then

$$\mathbf{x}' \cdot \mathbf{y}' = \mathbf{x} \cdot \mathbf{y},$$

i.e. the transformation R preserves the scalar product of all vectors. How do we know that other matrices R don't also achieve this? We shall now prove that *only* matrices satisfying $\mathsf{R}^\mathsf{T}\mathsf{R} = \mathtt{I}$ preserve the scalar product of all vectors.

First, let's forget momentarily that R is orthogonal, and simply assume that

$$\mathbf{x}' \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y},$$

for all $\mathbf{x}$ and $\mathbf{y}$, given some matrix transformation R. Using the arguments from the main lecture, the two sides of this equation can be rewritten as

$$\mathbf{x}^\mathsf{T}\mathsf{R}^\mathsf{T}\mathsf{R}\mathbf{y}' = \mathbf{x}^\mathsf{T}\mathbf{y},$$

which can be factorised as

$$\mathbf{x}^\mathsf{T}\left(\mathsf{R}^\mathsf{T}\mathsf{R} - \mathtt{I}\right)\mathbf{y} = 0,$$

i.e.

$$\mathbf{x} \cdot \left(\mathsf{R}^\mathsf{T}\mathsf{R} - \mathtt{I}\right)\mathbf{y} = 0.$$

Since we've assumed this must hold for all $\mathbf{x}$, it must certainly hold when $\mathbf{x} = \left(\mathsf{R}^\mathsf{T}\mathsf{R} - \mathtt{I}\right)\mathbf{y}$. Hence we need that

$$|\left(\mathsf{R}^\mathsf{T}\mathsf{R} - \mathtt{I}\right)\mathbf{y}|^2 = 0$$

for all $\mathbf{y}$. Since the modulus of a vector is only zero when the vector is zero, this means that

$$\left(\mathsf{R}^\mathsf{T}\mathsf{R} - \mathtt{I}\right)\mathbf{y} = \mathbf{0}$$

must be true for all $\mathbf{y}$. This in turn requires $\mathsf{R}^\mathsf{T}\mathsf{R} - \mathtt{I} = 0$, so $\mathsf{R}^\mathsf{T}\mathsf{R} = \mathtt{I}$ as desired.

You may well be asking how to prove the last step. Suppose that $\mathsf{A}\mathbf{y} = \mathbf{0}$ holds for all $\mathbf{y}$. It then certainly holds in the following case

$$\mathbf{y} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

But then, it is easy to confirm that

$$\mathsf{A}\mathbf{y} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \end{pmatrix} = \mathbf{0},$$

i.e. the first column of A must be all zeros. I'll leave it to you to complete the proof that $\mathsf{A} = 0$ if $\mathsf{A}\mathbf{y} = \mathbf{0}$ for all $\mathbf{y}$.