Orthogonal

First, an independent but important point: *You cannot make a meaningful assertion of the form*

$$\vec{x} = \vec{x}^{\|} + \vec{x}^{\perp}$$

without specifying the subspace to which the vectors on the right-hand side are parallel and perpendicular!

Projections are linear operators $T : \mathbb{R}^n \to \mathbb{R}^n$ satisfying $T^2 = T$. So if T is proj_W, do you know how to find W in terms of the usual things associated to T in linear algebra? How about $W^{\perp}$?

Let me give descriptions of five kinds of $n \times n$ matrices A (equivalently linear transformations $\mathbb{R}^n \to \mathbb{R}^n$):

1. $\|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$.
2. $A\vec{x} \cdot A\vec{y} = \vec{x} \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.
3. The column vectors of A comprise an orthonormal basis for $\mathbb{R}^n$.
4. $A^T A = I$ (so A is invertible with $A^{-1} = A^T$, and $AA^T = I$ as well).
5. The row vectors of A (transformed into column vectors in the standard way) comprise an orthonormal basis for $\mathbb{R}^n$.

The main point is that all five are equivalent, describing the orthogonal matrices. Mathematicians would set up the verification of that in such a way that one sees a cycle of statements, each implying the one ahead of it; that requires exactly five little verifications. I'll avoid this minimalism here.

<u>#1$\Longleftrightarrow$#2</u>: Length can be written in terms of dot product as $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$. This gives #2$\Longrightarrow$#1. On the other hand, there is the equality

$$\vec{x} \cdot \vec{y} = \tfrac{1}{2}[(\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) - (\vec{x} \cdot \vec{x}) - (\vec{y} \cdot \vec{y})],$$

which yields that #1$\Longrightarrow$#2.

<u>#2$\Longleftrightarrow$#3</u>: The i-th column vector of A is, as we know, $A\vec{e}_i$. Thus, if #2 holds, we get

$$A\vec{e}_i \cdot A\vec{e}_j = \vec{e}_i \cdot \vec{e}_j = \delta_{ij},$$

which gives #3. Conversely, we use #3 in expanding

$$\begin{aligned} A\vec{x} \cdot A\vec{y} &= A(\sum_{i=1}^{n} x_i \vec{e}_i) \cdot A(\sum_{j=1}^{n} y_j \vec{e}_j) \\ &= \sum_{i,j=1}^{n} x_i y_j (\vec{e}_i \cdot \vec{e}_j) = \sum_{i,j=1}^{n} x_i y_j \delta_{ij} \\ &= \sum_{i,j=1}^{n} x_i y_i = \vec{x} \cdot \vec{y}. \end{aligned}$$

In the above, we have used linearity in various ways.

#3$\Longleftrightarrow$#4: Think how matrix multiplication goes. The ij entry of the matrix product BA is the dot product of the i-th row vector of B and the j-th column vector of A. Said a little differently, it is dot product of the i-th column vector of B^T and the j-th column vector of A. When we put $B = A$, we get the matrix whose ij entry is the dot product of the i-th column vector of A and the j-th column vector of A, δ_{ij}, namely I. For square matrices, $A^T A = I \Leftrightarrow AA^T = I$. In the textbook, this can be found in Fact 2.4.9, whose justification is a bit buried.

#4$\Longleftrightarrow$#5: "$(A^T)^T A^T = I$" is equivalent to saying that the columns of A^T, i.e., the rows of A, are an orthonormal basis. $\square$

A little structure: Let's use the mathematician's symbol for the (more than just a) set of orthogonal matrices: $O(n)$. This is a subset of the full space of matrices, $L = \mathbb{R}^{n\times n}$. Now, L is a linear space of dimension n^2, but it also has multiplication of matrices as an additional operation, satisfying some algebraic properties (see Ch. 2.4). However, not every matrix has an inverse, which is a sort of "blemish." But $O(n)$ consists of invertible matrices, with A^{-1} given by the simple formula $A^{-1} = A^T$. Note that for arbitrary matrices, there is the formula $(AB)^T = B^T A^T$, parallel to the order-reversing formula for the inverse of a product of square matrices: $(AB)^{-1} = B^{-1}A^{-1}$. It follows that $O(n)$ is closed under matrix multiplication. However, $O(n)$ is *not* a linear space. (Why?)