LINEAR ALGEBRA

MATH 21B

ORTHOGONALITY

12.1. Orthogonality is an old concept in geometry. Already long before Pythagoras, examples like the 3-4-5 triangle were known to produce right angles. Pythagoras proved the general relation $a^2+b^2=c^2$ for the lengths of the vectors $\vec{v}, \vec{w}, \vec{v}-\vec{w}$ if $\vec{v}\cdot\vec{w}=0$. In linear algebra, it immediately follows from expanding $c^2=|\vec{v}-\vec{w}|^2=(\vec{v}-\vec{w})\cdot(\vec{v}-\vec{w})=\vec{v}\cdot\vec{v}+\vec{w}\cdot\vec{w}=|\vec{v}|^2+|\vec{w}|^2=a^2+b^2$, having removed the mixed terms $\vec{v}\cdot\vec{w}$ and $\vec{w}\cdot\vec{v}$.

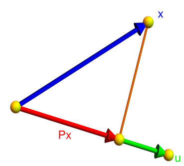

FIGURE 1. $P\vec{x}=\vec{u}(\vec{u}\cdot\vec{x})$ projects onto the line spanned by $\vec{u}$.

12.2. Remember that if $\vec{u}$ is a unit vector, then $\boxed{P\vec{x}=c\vec{u}}$ with $\boxed{c=\vec{x}\cdot\vec{u}}$ is the projection onto the line spanned by $\vec{u}$. We can also write this as $P\vec{x}=QQ^T\vec{x}$, where Q contains $\vec{u}$ as a column. For example, if $\vec{u}=[a,b]$, then

$$P=QQ^T=\begin{bmatrix} a \\ b \end{bmatrix}\begin{bmatrix} a & b \end{bmatrix}=\begin{bmatrix} a^2 & ba \\ ab & b^2 \end{bmatrix}.$$

12.3. To project onto a plane with orthogonal $\vec{u}_1, \vec{u}_2$ $P\vec{x}=c_1\vec{u}_1+c_2\vec{u}_2$, where $c_j=\vec{x}\cdot\vec{u}_j$. For example, in the standard basis, we have $P\vec{x}=x_1\vec{e}_1+\cdots+x_k\vec{e}_k$ where x_k are the standard coordinates of $\vec{x}$. We see that it is convenient to project, if we have orthogonal vectors in V of length 1. There is a name for that:

Definition: $\mathcal{B}=\{\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_k\}$ is called **orthonormal** if they are pairwise orthogonal $\vec{u}_l\cdot\vec{u}_k=0$ for $k\neq l$ and if they all have length 1.

Definition: A basis $\mathcal{B}$ is an **orthonormal basis** of V if it is a basis of V that is orthonormal.

12.4. Given an orthonormal basis $\mathcal{B}=(\vec{u}_1,\vec{u}_2,\ldots,\vec{u}_k)$ in V then every $\vec{v}$ in V can be written as $\vec{v}=c_1\vec{u}_1+\cdots+c_k\vec{u}_k$ with $c_j=\vec{v}\cdot\vec{u}_j$. For a general $\vec{v}\in V$ we have

Definition: $\text{proj}_V\vec{v}=c_1\vec{u}_1+\cdots+c_k\vec{u}_k$ with $c_j=\vec{v}\cdot\vec{u}_j$ is the **projection** onto V.

12.5. It is convenient to see this algebraically: let Q is the matrix which contains the vectors $\vec{u}_j$ as column vectors and let Q^T denote the matrix which contains the vectors $\vec{u}_j$ as row vectors. Lets compute Q^TQ:

$$Q = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} /2, \quad Q^TQ = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

Theorem: $\mathcal{B}$ is orthonormal if and only if Q^TQ is the identity matrix.

12.6. Just look what the vector multiplication does: every row of Q^T is one of the $\vec{u}_k$ and every column of Q is one of the $\vec{u}_l$. Now if you compute the product, you compute the dot product of $\vec{u}_k$ with $\vec{u}_l$ which is 0 if $k \neq l$ and equal to 1 if $k = l$.

12.7. We have seen that the projection onto the space spanned by an orthonormal basis is

$$P\vec{x} = c_1\vec{u}_1 + \cdots + c_k\vec{u}_k, c_k = \vec{x} \cdot \vec{u}_k .$$

We can rewrite this as $\boxed{P\vec{x} = Q^TQ\vec{x}}$ because $Q\vec{x} = \vec{c}$ and $Q\vec{c} = c_1\vec{u}_1 + \cdots + c_k\vec{u}_k$. In the future, we will derive a more general formula which works in the case when the vectors are no more orthonormal but where Q^TQ is still invertible. The **super cute formula** will be $P = Q(Q^TQ)^{-1}Q^T$. In the orthonormal case, where $Q^TQ = 1$ this simplifies to QQ^T.

12.8.

Theorem: Projection has the property that $P^2 = P$.

12.9. To verify this, just simplify $QQ^TQQ^T = Q(Q^TQ)Q^T$, using that $Q^TQ = 1$.

12.10. You can also check that $P\vec{x}$ and $\vec{x} - P\vec{x}$ are perpendicular: $(P\vec{x} \cdot (1 - P\vec{x}) = \vec{x}^T P(1-P)\vec{x} - 0$. The vector $P\vec{x}$ is sometimes denoted $\vec{x}^{\|}$ the vector perpendicular is written as $\vec{x}^{\perp}$. We have

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

12.11. The concept of orthogonality is also important in probability theory. If $X_1, \dots, X_n$ are centered random variables which are uncorrelated, then $\{X_1, \dots, X_n\}$ form an orthogonal system. If they have the standard deviation 1, this means that each of the variables has length 1, so that we deal with an orthonormal coordinate system. The probability theory of finitely many random variables is pretty much just linear algebra.