Lecture 17: Orthogonality

Two vectors $\vec{v}$ and $\vec{w}$ are called **orthogonal** if their dot product is zero $\vec{v} \cdot \vec{w} = 0$.

1 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 6 \\ -3 \end{bmatrix}$ are orthogonal in $\mathbf{R}^2$.

2 $\vec{v}$ and $\vec{w}$ are both orthogonal to the cross product $\vec{v} \times \vec{w}$ in $\mathbf{R}^3$. The dot product between $\vec{v}$ and $\vec{v} \times \vec{w}$ is the determinant
$$\det\left(\begin{bmatrix} v_1 & v_2 & v_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}\right) .$$

$\vec{v}$ is called a **unit vector** if its length is one: $||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}} = 1$.

A set of vectors $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is called **orthogonal** if they are pairwise orthogonal. They are called **orthonormal** if they are also unit vectors. A basis is called an **orthonormal basis** if it is a basis which is orthonormal. For an orthonormal basis, the matrix with entries $A_{ij} = \vec{v}_i \cdot \vec{v}_j$ is the unit matrix.

Orthogonal vectors are linearly independent. A set of n orthogonal vectors in $\mathbf{R}^n$ automatically form a basis.

Proof: The dot product of a **linear relation** $a_1\vec{v}_1 + \ldots + a_n\vec{v}_n = 0$ with $\vec{v}_k$ gives $a_k \vec{v}_k \cdot \vec{v}_k = a_k ||\vec{v}_k||^2 = 0$ so that $a_k = 0$. If we have n linear independent vectors in $\mathbf{R}^n$, they automatically span the space because the fundamental theorem of linear algebra shows that the image has then dimension n.

A vector $\vec{w} \in \mathbf{R}^n$ is called **orthogonal** to a linear space V, if $\vec{w}$ is orthogonal to every vector $\vec{v} \in V$. The **orthogonal complement** of a linear space V is the set W of all vectors which are orthogonal to V.

The orthogonal complement of a linear space V is a linear space. It is the kernel of A^T, if the image of A is V.

To check this, take two vectors in the orthogonal complement. They satisfy $\vec{v} \cdot \vec{w}_1 = 0, \vec{v} \cdot \vec{w}_2 = 0$. Therefore, also $\vec{v} \cdot (\vec{w}_1 + \vec{w}_2) = 0$.

Pythagoras theorem: If $\vec{x}$ and $\vec{y}$ are orthogonal, then $||\vec{x} + \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2$.

Proof. Expand $(\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y})$.

Cauchy-Schwarz: $|\vec{x} \cdot \vec{y}| \leq ||\vec{x}|| \, ||\vec{y}||$.

Proof: $\vec{x} \cdot \vec{y} = ||\vec{x}||||\vec{y}|| \cos(\alpha)$. If $|\vec{x} \cdot \vec{y}| = ||\vec{x}||||\vec{y}||$, then $\vec{x}$ and $\vec{y}$ are parallel.

Triangle inequality: $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof: $(\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) = ||\vec{x}||^2 + ||\vec{y}||^2 + 2\vec{x} \cdot \vec{y} \leq ||\vec{x}||^2 + ||\vec{y}||^2 + 2||\vec{x}||||\vec{y}|| = (||\vec{x}|| + ||\vec{y}||)^2$.

Angle: The **angle** between two vectors $\vec{x}, \vec{y}$ is $\alpha = \arccos\left(\frac{\vec{x} \cdot \vec{y}}{||\vec{x}||||\vec{y}||}\right)$.

$\cos(\alpha) = \frac{\vec{x} \cdot \vec{y}}{||\vec{x}||||\vec{y}||} \in [-1, 1]$ is the **statistical correlation** of $\vec{x}$ and $\vec{y}$ if the vectors $\vec{x}, \vec{y}$ represent data of zero mean.

3 Express the fact that $\vec{x}$ is in the kernel of a matrix A using orthogonality. **Answer** $A\vec{x} = 0$ means that $\vec{w}_k \cdot \vec{x} = 0$ for every row vector $\vec{w}_k$ of $\mathbf{R}^n$. Therefore, the orthogonal complement of the row space is the kernel of a matrix.

The **transpose** of a matrix A is the matrix $(A^T)_{ij} = A_{ji}$. If A is a $n \times m$ matrix, then A^T is a $m \times n$ matrix. Its rows are the columns of A. For square matrices, the transposed matrix is obtained by reflecting the matrix at the diagonal.

4 The transpose of a vector $A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is the row vector $A^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$.

The transpose of the matrix $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is the matrix $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.

$(AB)^T = B^T A^T$, $(A^T)^T = A$, $(A^{-1})^T = (A^T)^{-1}$.
$v^T w$ is the dot product $\vec{v} \cdot \vec{w}$.
$\vec{x} \cdot A\vec{y} = A^T\vec{x} \cdot \vec{y}$.

The proofs are direct computations. Here is the first identity:
$$(AB)^T_{kl} = (AB)_{lk} = \sum_i A_{li}B_{ik} = \sum_i B^T_{ki}A^T_{il} = (B^T A^T)_{kl} .$$

A linear transformation is called **orthogonal** if $A^T A = I_n$.

We see that a matrix is orthogonal if and only if the column vectors form an orthonormal basis.

Orthogonal matrices preserve length and angles. They satisfy $A^{-1} = A^T$.

5 A rotation is orthogonal.

6 A reflection is orthogonal.

Orthogonality over time

- From -2800 BC until -2300 BC, Egyptians used ropes divided into length ratios like $3 : 4 : 5$ to build triangles. This allowed them to triangulate areas quite precisely: for example to build irrigation needed because the Nile was reshaping the land constantly or to build the pyramids: for the **great pyramid at Giza** with a base length of 230 meters, the average error on each side is less then 20cm, an error of less then $1/1000$. A key to achieve this was **orthogonality**.

- During one of Thales (-624 BC to (-548 BC)) journeys to Egypt, he used a geometrical trick to **measure the height** of the great pyramid. He measured the size of the shadow of the pyramid. Using a stick, he found the relation between the length of the stick and the length of its shadow. The same length ratio applies to the pyramid (**orthogonal** triangles). Thales found also that triangles inscribed into a circle and having as the base as the diameter must have a right angle.

- The Pythagoreans (-572 until -507) were interested in the discovery that the squares of a lengths of a triangle with two **orthogonal** sides would add up as $a^2 + b^2 = c^2$. They were puzzled in assigning a length to the diagonal of the unit square, which is $\sqrt{2}$. This number is irrational because $\sqrt{2} = p/q$ would imply that $q^2 = 2p^2$. While the prime factorization of q^2 contains an even power of 2, the prime factorization of $2p^2$ contains an odd power of 2.

- Eratosthenes (-274 until 194) realized that while the sun rays were **orthogonal** to the ground in the town of Scene, this did no more do so at the town of Alexandria, where they would hit the ground at 7.2 degrees). Because the distance was about 500 miles and 7.2 is 1/50 of 360 degrees, he measured the circumference of the earth as 25'000 miles - pretty close to the actual value 24'874 miles.

- Closely related to **orthogonality** is **parallelism**. Mathematicians tried for ages to prove Euclid's parallel axiom using other postulates of Euclid (-325 until -265). These attempts had to fail because there are geometries in which parallel lines always meet (like on the sphere) or geometries, where parallel lines never meet (the Poincaré half plane). Also these geometries can be studied using linear algebra. The geometry on the sphere with **rotations**, the geometry on the half plane uses Möbius transformations, 2×2 matrices with determinant one.

- The question whether the angles of a right triangle do always add up to 180 degrees became an issue when geometries where discovered, in which the measurement depends on the position in space. Riemannian geometry, founded 150 years ago, is the foundation of **general relativity**, a theory which describes gravity geometrically: the presence of mass bends space-time, where the dot product can depend on space. **Orthogonality** becomes relative. On a sphere for example, the three angles of a triangle are bigger than 180^+. Space is curved.

- In **probability theory**, the notion of **independence** or **decorrelation** is used. For example, when throwing a dice, the number shown by the first dice is independent and decorrelated from the number shown by the second dice. Decorrelation is identical to **orthogonality**, when vectors are associated to the random variables. The **correlation coefficient** between two vectors $\vec{v}, \vec{w}$ is defined as $\vec{v} \cdot \vec{w}/(|\vec{v}||\vec{w}|)$. It is the cosine of the angle between these vectors.

- In **quantum mechanics**, states of atoms are described by functions in a linear space of functions. The states with energy $-E_B/n^2$ (where $E_B = 13.6eV$ is the Bohr energy) in a hydrogen atom. States in an atom are **orthogonal**. Two states of two different atoms which don't interact are **orthogonal**. One of the challenges in quantum computation, where the computation deals with qubits (=vectors) is that orthogonality is not preserved during the computation (because we don't know all the information). Different states can interact.

Homework due March 9, 2011

1 a) Assume X, Y, Z, U, V are independent random variables which have all standard deviation 1. Find the standard deviation of $X + Y + 2Z + 3U - V$.
b) We have two random variables X and Y of standard deviation 1 and 2 and correlation -0.5. Can you find a combination $Z = aX + bY$ such that X, Z are uncorrelated?
c) Verify that the standard deviation of $X + Y$ is smaller or equal than the sum of the standard deviations of X and Y.

2 a) Verify that if A, B are orthogonal matrices then their product $A.B$ and $B.A$ are orthogonal matrices.
b) Verify that if A, B are orthogonal matrices, then their inverse is an orthogonal matrix.
c) Verify that 1_n is an orthogonal matrix.
These properties show that the space of $n \times n$ orthogonal matrices form a "group". It is called $O(n)$.

3 Given a basis $\mathcal{B}$, we describe a process called **Gram-Schmidt orthogonalization** which produces an orthonormal basis.
If $\vec{v}_1, \ldots, \vec{v}_n$ are the basis vectors let $\vec{w}_1 = \vec{v}_1$ and $\vec{u}_1 = \vec{w}_1/||\vec{w}_1||$. The Gram-Schmidt process recursively constructs from the already constructed orthonormal set $\vec{u}_1, \ldots, \vec{u}_{i-1}$ which spans a linear space V_{i-1} the new vector $\vec{w}_i = (\vec{v}_i - \mathrm{proj}_{V_{i-1}}(\vec{v}_i))$ which is orthogonal to V_{i-1}, and then normalizing $\vec{w}_i$ to to get $\vec{u}_i = \vec{w}_i/||\vec{w}_i||$. Each vector $\vec{w}_i$ is orthonormal to the linear space V_{i-1}. The vectors $\{\vec{u}_1, .., \vec{u}_n\}$ form an orthonormal basis in V.
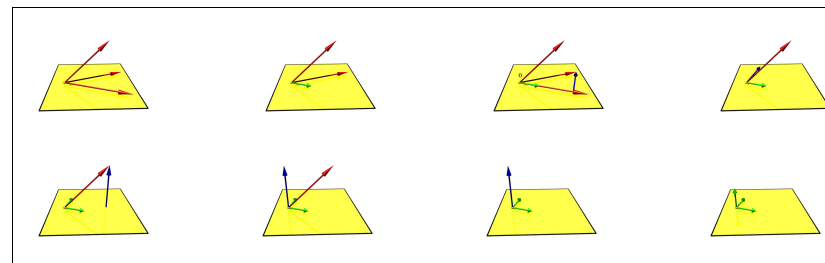
Find an orthonormal basis for $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$.