CS/ENGRI 172, Fall 2003: Computation, Information, and Intelligence
9/22/03: Vector Notation for Function Input

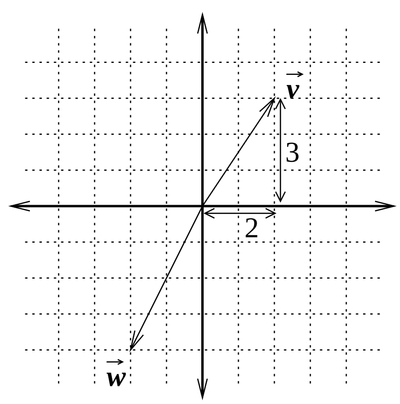
Defining Feature Vectors

In this course, our view of machine learning will be that of learning a function, which is actually a quite general framework. We will consider the input to the functions to be learned to be (numerically-valued, finite) *feature vectors*, or ordered tuples. An example of our notation is as follows, indicating that $\vec{x}$ is an n-dimensional vector with its n *components* being x_1 through x_n:

$$\vec{x} = (x_1,\, x_2,\, \ldots,\, x_n).$$

It is standard to indicate vectors with an arrow or boldfacing, to distinguish them from components, which are numbers ("scalars").

Two-dimensional vectors are probably the most familiar to you, since they can be associated with points in the plane. We can represent a two-dimensional vector $\vec{x}$ as a directed arrow from the origin to the point specified by the ordered pair $(x_1,\, x_2)$. So, for example, if $\vec{v} = (2,\, 3)$ and $\vec{w} =$ (-1, -4), then we have the following picture:



This association of vectors with directed arrows from the origin to the point specified by the components can be continued into higher dimensions, though it becomes more difficult to draw.

A confusing fact of life is that it is common to use the letters "x" and "y" to indicate both vector names, vector components, and, in the case of two-dimensional vectors, coordinates; for example, one might say "the x-coordinate of the vector $\vec{x} = (x_1,\, x_2)$ is x_1". We will try to avoid such confusing utterances, but be aware of the potential for confusion and watch notation and context carefully.

Basic Vector Manipulation

Vectors are mathematically manipulated in a *component-wise* manner, so that we will, in general, perform an operation on a vector by performing a related operation or operations on its component scalars. For manipulations between two vectors, we will require that both vectors have the same dimensionality.

Vector *addition* between two vectors $\vec{x} = (x_1, x_2, \ldots, x_n)$ and $\vec{y} = (y_1, y_2, \ldots, y_n)$ is defined as:

$$\vec{x} + \vec{y} \stackrel{\text{def}}{=} (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

The result of vector addition is another vector of the same dimensionality. For example, if $\vec{x}$ = (2, 3) and $\vec{y}$ = (-1, -4), $\vec{x} + \vec{y} = (1, -1)$.

The *inner product* (or *dot product*) between two vectors of the same dimensionality $\vec{x} = (x_1, x_2, \ldots, x_n)$ and $\vec{y} = (y_1, y_2, \ldots, y_n)$ is defined as:

$$\vec{x} \cdot \vec{y} \stackrel{\text{def}}{=} \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

The result of the inner product of two vectors is a *scalar*. For example, if $\vec{x}$ = (2, 3) and $\vec{y}$ = (-1, -4), $\vec{x} \cdot \vec{y}$ = -2 - 15 = -17.

The *length* of a vector $\vec{x} = (x_1, x_2, \ldots, x_n)$ can be computed via the inner product:

$$\text{length}(\vec{x}) = \sqrt{\vec{x} \cdot \vec{x}}$$

For example, consider the two-dimensional case: $\text{length}((x_1, x_2)) = \sqrt{(v_1, v_2) \cdot (v_1, v_2)} = \sqrt{v_1^2 + v_2^2}$, which is exactly the Pythagorean theorem. The length of a vector is also a scalar.

An extremely handy fact is the following identity:

$$\begin{aligned} \vec{x} \cdot \vec{y} &= \text{length}(\vec{x})\text{length}(\vec{y})\cos(\angle(\vec{x}, \vec{y})) \\ &= \text{length}(\vec{y})\text{proj}(\vec{x}, \vec{y}) \end{aligned}$$

where for notational convenience we consider the projection of one vector onto another to be a length, i.e. a scalar, rather than a vector. To convince yourself, recall that "cosine = adjacent over hypotenuse", and consider the diagram: