Oftentimes, a linear system $A\mathbf{x} = \mathbf{b}$ that arises from experimental data is inconsistent. However, an acceptable substitute for a solution, is a vector $\hat{\mathbf{x}}$ that makes the distance between $A\hat{\mathbf{x}}$ and $\mathbf{b}$ as small as possible. In statistics, for example, the $\hat{\mathbf{x}}$ vector is camouflaged as the least squares solution (examined in a later section) of $A\mathbf{x} = \mathbf{b}$ which minimizes the distance (sum) of squared errors.

Before looking at the least square solution of $A\mathbf{x} = \mathbf{b}$, we have to formally define the concept of inner product which will lead us to the definition of distance.

The Inner Product

Suppose that $\mathbf{v}$ and $\mathbf{w}$ are two vectors in $\Re^n$. The *single* number $\mathbf{v}^T\mathbf{w}$ is called the **inner product** of $\mathbf{v}$ and $\mathbf{w}$, and it is often written as $<\mathbf{v}, \mathbf{w}>$ or as $\mathbf{v} \cdot \mathbf{w}$ (also called the **dot product**).

For example,

$$\text{if } \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix}$$

then the inner product of $\mathbf{v}$ and $\mathbf{w}$, is $v_1w_1 + v_2w_2 + \cdots + v_nw_n$.

Example

Compute $<\mathbf{v}, \mathbf{w}>$ and $<\mathbf{w}, \mathbf{v}>$ when $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$

$$<\mathbf{v}, \mathbf{w}> = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} = (v_1)(w_1) + (v_2)(w_2) + (v_3)(w_3)$$

$$<\mathbf{w}, \mathbf{v}> = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = (w_1)(v_1) + (w_2)(v_2) + (w_3)(v_3)$$

It is clear from calculations that $<\mathbf{v}, \mathbf{w}>$ and $<\mathbf{w}, \mathbf{v}>$ are equal for any $\mathbf{v}, \mathbf{w} \in \Re^n$.