COMMENTS ON THE DOT PRODUCT

1. ON COLUMN VECTORS

The book uses (a, b, c) interchangably with $\begin{bmatrix} a & b & c \end{bmatrix}$. So we are stuck with

$$(a, b, c)^T = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

The *dot product*, less commonly called the *scalar product*, is defined on a pair of vectors in $\mathbb{R}^n$ as follows:

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

or

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

If you are willing (as our author is) to pretend a scalar is the same as a one-by-one matrix, then

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}.$$

The works because

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

What I dislike about the book's notation is you will run into this in the problems: if

$$\mathbf{x} = (a, b, c)^T \text{ and } \mathbf{y} = (r, s, t)^T$$

then

$$\mathbf{x}^T \mathbf{y} \quad = \quad \left((a, b, c)^T\right)^T (r, s, t)^T.$$

This works, since

$$\mathbf{x}^T \mathbf{y} = [ar + bs + ct] = ar + bs + ct$$

but you will find it easier to just make this replacement in this chapter

$$\mathbf{x}^T \mathbf{y} \rightsquigarrow \mathbf{x} \cdot \mathbf{y}$$

as in these examples:

The length becomes

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

The cosine of the angle between vectors becomes

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

2. On n-tuples

If you are working directly with ordered n-tuples, the dot product is just

$$(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

3. On Row Vectors

You will see the dot product of two row vectors:

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

In this, case, if you approve of treating a 1-by-1 matrix as a scalar, this holds:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}\mathbf{y}^T.$$

4. Why 1-by-1 isn't scalar

If we strip off the braces from all one-by-one matrices, then odd things happen. For example

$$\left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

and yet

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) = !?$$

is not even defined. So we've lost the associative law.

5. Dealing with Matlab

The 1-by-1 is scalar nonsense breaks a few things in Matlab. As in the last section:

```
( [1 1] * [1;1] ) * [1 1;1 1]
```

returns a 2-by 2 matrix, while

```
[1 1] * ( [1;1] * [1 1;1 1] )
```

returns an error.

In any case, if you want the dot product of two column vectors in Matlab, you need to use `x'*y`. If you want the dot product of two row vectors you use `x*y'`.

6. Some Axioms

Using what we know about transpose and matrix multiplication, you can figure out

$$(\mathbf{x}_1 + \mathbf{x}_2)^T \mathbf{y} = (\mathbf{x}_1^T + \mathbf{x}_2^T)\mathbf{y} = \mathbf{x}_1^T \mathbf{y} + \mathbf{x}_2^T \mathbf{y}$$

but it is a good idea to see this directly in terms of the dot product.

Here are some formulas for the dot product:

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$$
$$(\alpha \mathbf{x}) \cdot \mathbf{z} = \alpha(\mathbf{x} \cdot \mathbf{z})$$
$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$$
$$\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$$
$$\mathbf{x} \cdot (\alpha \mathbf{y}) = \alpha \mathbf{x} \cdot \mathbf{y}$$
$$\mathbf{x} \cdot \mathbf{y} = 0 \quad \iff \mathbf{x} \perp \mathbf{y} \quad \text{(this is a definition)}$$