Sec. 3.2 Inequalities

I am going to do this slightly differently from the book.
Lets first talk about projections onto a line.

Projections onto lines
Please refer to the video for the sketch.

The shortest distance from a point $\vec{b}$ onto a line through $\vec{a}$ is via a line through $\vec{b}$ that is perpendicular to the line through $\vec{a}$. This will meet the line through $\vec{a}$ at a point $\vec{p}$. Then $\vec{p}$ is simply a scaled version of $\vec{a}$, so $\vec{p} = \hat{x}\vec{a}$, and the perpendicular line is $\vec{b} - \vec{p}$. Then $\vec{a}^T(b - \hat{x}\vec{a}) = 0$, and solving for $\hat{x}$ give us $\hat{x} = (\vec{a}^T\vec{b})/(\vec{a}^T\vec{a})$.

Definition 1. The projection of the vector $\vec{b}$ onto the line in the direction of $\vec{a}$ is

$$\vec{p} = \hat{x}\vec{a} = \left(\frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}}\right)\vec{a}. \tag{1}$$

We will use projections for least squares solutions to singular systems, and to derive the Cauchy-Schwartz inequality. Before we go into that, lets first do some examples with projections.

Ex: Project $b = (1,2,3)$ onto the line through $a = (1,1,1)$ to get $\hat{x}$ and p.

Solution:

$$\hat{x} = \frac{a^Tb}{a^Ta} = \frac{6}{3} = 2 \Rightarrow p = \hat{x}a = (2,2,2).$$

Ex: Consider the vectors

$$\vec{u} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}, \qquad \vec{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

(1) For u onto v we do

$$\hat{x} = \frac{v^Tu}{v^Tv} = \frac{8}{2} = 4 \Rightarrow p = \hat{x}v = (4,-4,0)$$

(2) For v onto u we do

$$\hat{x} = \frac{u^Tv}{u^Tu} = \frac{8}{35} \Rightarrow p = \hat{x}u = \left(\frac{8}{7}, -\frac{24}{35}, \frac{8}{35}\right)$$

The Cauchy-Schwartz inequality

As mentioned before, this leads us to perhaps the most important inequality in mathematics. Lets derive it here.

Recall that the error vector $e = b - p \Rightarrow \|e\| = \|b - p\|$, then writing out the equation for the projection vector and squaring gives us

$$\left\|b - \frac{a^Tb}{a^Ta}a\right\|^2 = \left(b - \frac{a^Tb}{a^Ta}a\right)^T\left(b - \frac{a^Tb}{a^Ta}a\right) = b^Tb - 2\frac{(a^Tb)^2}{a^Ta} + \left(\frac{a^Tb}{a^Ta}\right)^2 a^Ta = \frac{(b^Tb)(a^Ta) - (a^Tb)^2}{a^Ta} \geq 0$$

since we squared the left hand side. Since the denominator is finite,

$$(b^Tb)(a^Ta) - (a^Tb)^2 \geq 0 \Rightarrow (a^Tb)^2 \leq \|a\|^2\|b\|^2.$$

Another way we can derive this inequality is through the law of cosines.

$$\left|\frac{a^Tb}{\|a\|\|b\|}\right| = |\cos\theta| \leq 1 \Rightarrow |a^Tb| \leq \|a\|\|b\|.$$

Theorem 1. *All inner products $\langle a, b\rangle$ satisfy the Cauchy-Schwartz inequality*

$$\left|a^Tb\right| = \|a\|\|b\|. \tag{2}$$