Gram-Schmidt

Alright guys, here is how we want to do is, given linearly independent xectors $\{\vec{x_1}, \ldots, \vec{x_n}\}$, find a set of orthgonal vectors $\{\vec{v_1}, \ldots, \vec{v_n}\}$ such that $Span(\{\vec{x_1}, \ldots, \vec{x_n}\}) = Span(\{\vec{v_1}, \ldots, \vec{v_n}\})$. Note that the set $\{\vec{v_1}, \ldots, \vec{v_n}\}$ will be orthogonal and not orthonomal. To make the set orthonormal, just take $\{\vec{u_1}, \ldots, \vec{u_n}\}$, where $\vec{u_i} = \frac{\vec{v_i}}{||\vec{v_i}||}$, for $1 \leq i \leq n$. From here on out, I'll just say v_i instead of $\vec{v_i}$.
So, we set out to get our v_i in terms of our x_i.

Set $v_1 = x_1$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1$$

...

$$v_n = x_n - \frac{x_n \cdot v_{n-1}}{v_{n-1} \cdot v_{n-1}} - \ldots - \frac{x_n \cdot v_2}{v_2 \cdot v_2} v_2 - \frac{x_n \cdot v_1}{v_1 \cdot v_1} v_1$$

Again, to make the v_i orthonormal, just take $frac\vec{v_i}||\vec{v_i}||$.
I'm not going to explain here why $Span(\{\vec{x_1}, \ldots, \vec{x_n}\}) = Span(\{\vec{v_1}, \ldots, \vec{v_n}\})$ (although you should be able to figure it out), or why the v_i are orthogonal- just know how to do this!

QR Factorization

The claim here is given any m by n matrix M, we can find an m by n matrix Q and an n by n matrix R such that R is upper triangular (all entries below the diagonal are zero) and all the entries on the diagonal are positive.
Take our matrix M as column vectors $[x_1 : x_2 : \ldots : x_n]$. Then we set $Q = [v_1 : v_2 : \ldots : v_n]$, where we got the v_i by using our Gram-Schmidt process.
Realize from above that we have:

$$v_1 = x_1$$

$$x_2 = v_2 + \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$x_3 = v_3 + \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 + \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1$$

...

$$x_n = v_n + \frac{x_n \cdot v_{n-1}}{v_{n-1} \cdot v_{n-1}} + \ldots + \frac{x_n \cdot v_2}{v_2 \cdot v_2} v_2 + \frac{x_n \cdot v_1}{v_1 \cdot v_1} v_1$$

Realize, then, that the matrix R is just the coefficients for the linear combinations above! So, in particular, $R_{11} = \ldots = R_{nn} = 1 > 0$, and if R_{ij} is above the diagonal, $R_{ij} = \frac{x_j \cdot v_i}{v_i \cdot v_i}$.
Note! This all changes if we require that the v_i are orthonormal, in particular, in the above equations, we'll have to divide the ith equation by $||v_i||$, but the idea is the same.
Hope this helps.
Good luck!