The Cholesky Factorization

Symmetric matrices are important because they are common in applications, have some very nice properties, and because the symmetry can be exploited by algorithms to save time and memory. For example, we know that if $A = A^t$ has an LU factorization, then $A = LDL^t$ can be computed in about $n^3/3$ flops.

Because of small pivots, the LDL^t algorithm is not backward stable for general symmetric matrices. However, there are matrices for which pivoting is never needed: for which the diagonal element is always the |largest| element in its column. The most important of these are the *symmetric positive definite* (spd) matrices.

$A \in \mathbb{R}^{n\times n}$ is spd if it is symmetric and if for all $x \neq 0 \in \mathbb{R}^n$, $x^tAx > 0$. All matrices of the form X^tX are spd iff X has linearly independent columns.

A matrix has an LDL^t factorization with $d_{ii} > 0$ iff it is spd. In that case we may write

$$A = LDL^t = LD^{1/2}D^{1/2}L^t \equiv GG^t.$$

The *Cholesky factorization*, which computes $A = GG^t$ directly, is a simple and popular alternative to the LDL^t factorization for spd matrices.

The classical algorithm appears easily by looking at the columns of $A = GG^t$ with lower triangular G. Assume we know the first $k-1$ columns of G, and look at the k^{th} column of $A = GG^t$:

$$a_k \equiv Ae_k = GG^te_k = [g_1, g_2, \ldots, g_n]z.$$

Here z^t is the k^{th} row of (lower triangular) G, so we can write

$$a_k = \sum_{i=1}^{k} g_{ki}g_i,$$

or

$$g_{kk}g_k = a_k - \sum_{i=1}^{k-1} g_{ki}g_i.$$

In particular $g_{kk}^2 = a_{kk} - \sum_{j=1}^{k-1} g_{kj}^2$, so we take the positive root and solve for g_k:

$$g_k = (a_k - \sum_{i=1}^{k-1} g_{ki}g_i)/g_{kk}.$$

This method runs to completion (no zero or complex roots) iff A is spd. It requires $\frac{1}{3}n^3 + \mathrm{O}(n^2)$ flops, and no extra memory if the lower triangular of A is overwritten by that of G. With respect to rounding errors, the computed $\tilde{G}$ satisfies

$$\tilde{G}\tilde{G}^t = A + \delta A, \text{ where } \|\delta A\| \leq 12n^2\mu\|A\|.$$