Symmetric Indefinite Factorization

Recall $A = A^T$ is indefinite.

If there are nonzero vectors $x, y$ such that

$x^T Ax > 0$ and $y^T Ay < 0$.

Examples:

a) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
   Take $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $y = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ above.

b) $A = \begin{pmatrix} \varepsilon & 1 \\ 1 & \varepsilon \end{pmatrix}$, $0 < |\varepsilon| < 1$
   Take $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $y = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ above.

Fact: No symmetric permutation of $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has an LDLT factorization
   with $L$ unit lower triangular and $D$ diagonal.

Why? $PAP^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ for $P = I_2$ or $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
   (which are the only possibilities).
One might think that the basic step
\[ (x \ a^T) = (I \ a) \begin{pmatrix} a & -\alpha aI \end{pmatrix} (0 \ I) \]
with \( d = a \cdot \frac{1}{2} \) could be done regardless of the sign of \( \alpha \).

However, this fails completely if \( \alpha = 0 \) and in the case
\[ A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad |x| < 1 \]
the factorization is
\[ A = \begin{pmatrix} 1 & 0 \\ 0 & e^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & 1 \end{pmatrix} \]
\[ = L \ D \ L^T \]

Note: \( L(2,1) \to \infty \) as \( |x| \to 0 \)
\( D(2,2) \to -\infty \) " "

So the nice boundedness property
of sparse Cholesky factorization is lost.
Such a factorization is unstable numerically.
The remedy is the Bunch-Parlett Bunch-Kaufman algorithms (The first uses complete pivoting, the second is a partial pivoting variant).

Diagram of Bunch-Kaufman strategy

Denote the elements of $A$ by $a_{ij}$

\[ j = \text{index of max element } |a_{i1}| \]

\[ \text{i.e. } |a_{ij}| = \| A_{ij} \|_\infty \]

\[ = \arg \max_2^\infty |a_{i1}| i^3 \]

\[ \Theta = 0.64 \quad (\approx \frac{1+\sqrt{17}}{8} - \text{magic number}) \]

\[ \{ \begin{array}{l}
\varepsilon \geq \max \{ \max_j \sum_{k \neq j} |a_{k\ell}| \} \\
\{a \text{ pivoting "threshold"}\}
\end{array} \]

\[ m = \max \{ \sum_{j \neq i} |a_{ij}| \} \]

\[ = \max \{ \max_{j \neq i} \sum_{k \neq j} |a_{kj}| \} \]

(lower triangle ref. only)
The 2×2 matrix \( \begin{pmatrix} a_{ij} & x_{ij} \\ x_{ij} & a_{ij} \end{pmatrix} \)

is analyzed using the threshold \( \Theta \) and the max element \( a_{ij} \) (in the \( j \)th column) to decide the pivoting strategy.

*In words:*

i) If \( |a_{21}| \) is "large enough" use it as a 1×1 pivot (no pivoting required).

ii) Else, if \( |a_{ij}| \) is "large enough"

interchange row \( j \) with row 1

\( \{ \text{col } j \text{ with col 1} \) and use \( a_{ij} \) (now in the (1,1) position) as a 1×1 pivot.

iii) Else interchange row 2 with row \( j \), col 2 with col \( j \) and take

\( \begin{pmatrix} x_{11} & x_{1j} \\ x_{2j} & a_{2j} \end{pmatrix} \)

as a 2×2 pivot.
The basic factorization step in cases (i) and (ii) is the same as for Cholesky, even if a negative entry is placed in the (1,1) position after pivoting.

In case (iii), let \( D = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \)

Let 
\[
P_1 A P_1^T = \begin{bmatrix} D_1 & A_{21}^T \\ A_{21} & A_{22} \end{bmatrix} = L_1 \begin{bmatrix} I_2 & 0 \\ L_{21} & T \end{bmatrix} \begin{bmatrix} D_1 & 0 \\ 0 & A_{22} - A_{21} D_1 A_{21}^T \end{bmatrix} \begin{bmatrix} I_2 & L_{21}^T \\ 0 & I \end{bmatrix}
\]

Where 
\( A_{21} = A(3:n, 1:2) \)
\( A_{22} = A(3:n, 3:n) \)

Then the basic 2x2 factorization step is

Now, continue applying pivot and factorization to
\( A_{22} - A_{21} D_1 A_{21}^T \)

Note: \( D_1^{-1} = \frac{1}{\Delta} \begin{bmatrix} \alpha_{22} & -\alpha_{21} \\ -\alpha_{21} & \alpha_{11} \end{bmatrix}, \Delta = \alpha_{11} \alpha_{22} - \alpha_{21}^2 \)
Upon completion, the factorization will be

\[ \mathbf{R} \mathbf{A} \mathbf{P}^T = \mathbf{L} \mathbf{D} \mathbf{L}^T \]

\( \mathbf{D} \) is block-diagonal with 2x2 and 1x1 diagonal blocks.

\[ \mathbf{D} = \begin{bmatrix} \mathbf{D}_{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2} \end{bmatrix} \]

\( \mathbf{L} \) is unit lower triangular (1's on diagonal) and if

\[ \mathbf{D}(i:i+1, j:i+1) \ is \ 2 \times 2 \]

Then

\[ \mathbf{L}(i:i+1, j:i+1) = \begin{bmatrix} 1 & 0 \\ \mathbf{0} & \mathbf{L} \end{bmatrix} \]

Only the strict lower triangle of \( \mathbf{L} \) is stored and it overwritten the
strict lower triangle of \( \mathbf{A} \) (as in LDL' factorization).
Def: Inertia $(A) = (n, z, p)$

\[ n = \# \text{ negative eigenvalues} \]
\[ z = \# \text{ zero eigenvalues} \]
\[ p = \# \text{ positive eigenvalues} \]

Fact: Inertia $(D) = \text{Inertia} (A)$

$\text{Inertia} (D) = \# \text{2x2 blocks} + \# \text{ neg. 1x1 blocks}$

Each 2x2 block will have a pos. +/- neg. eigenvalue.

Why? Sylvester's law of inertia (Sylvester's inertia theorem).

Thus: If $A = A^T$ and $M$ is nonsingular

Then: Inertia $(A) = \text{Inertia} (MAM^T)$.

$MAM^T$ (M non-sing) is called a congruence of $A$. 
Formally, the pivoting decision ($\theta = 0.64$) is

$$j = \arg \min_{i} |a_{i,j}|$$

$$\mu = \max_{j} |a_{j,j}|$$

$$\text{if } (|a_{1,1}| \geq \theta |a_{j,1}| \text{ or } |a_{1,1}| \geq \theta |a_{j,1}|^2)$$

$$p_1 = 1 \quad (\text{no pivot})$$

else

$$\text{if } (|a_{j,j}| \geq \theta \mu)$$

$$p_1 = j \quad (\text{interchange 1 with } j)$$

$$p_2 = 1 \quad (1 \times 1 \text{ pivot})$$

else

$$p_1 = j \quad (\text{interchange 2 with } j)$$

$$p_2 = j \quad (2 \times 2 \text{ pivot})$$

end if

end if
Note: If a $2 \times 2$ pivot is done then

(i) $|\alpha_{jj}| < \Theta | \alpha_{jj} |$ and (ii) $| |\alpha_{jj}| < \Theta | \alpha_{jj} |^2$

and (iii) $| \alpha_{jj} | < \Theta n$

Thus

$\Lambda = \det \left( \begin{array}{cc} \alpha_{ii} & \alpha_{ii} \\ \alpha_{ij} & \alpha_{jj} \end{array} \right) = \alpha_{ii} \alpha_{jj} - \alpha_{ij}^2$

satisfies

$\Lambda < |\alpha_{ii}| |\alpha_{jj}| - (\alpha_{ij})^2$

$< | \alpha_{ii} | n \Theta - |\alpha_{jj}|^2$ (due to (iii))

$< \Theta^2 |\alpha_{jj}|^2 - |\alpha_{jj}|^2$ (due to (ii))

$= - (1 - \Theta^2) |\alpha_{jj}|^2 < 0$

Hence $D = \left( \begin{array}{cc} \alpha_{ii} & \alpha_{ii} \\ \alpha_{ij} & \alpha_{jj} \end{array} \right)$ has a positive and a negative eigenvalue.

$\Delta = \lambda_{1,2}^2$, where $\lambda_{1,2}$ are the eigenvalues of $D$. 
Note also that if \( A \) is spd,
then
\[
(x_i^2 \quad x_j) \text{ is also s.p.d.}
\]

Hence \( x_i^2 x_j - x_j x_i > 0 \).

This
\[
1 \mu_x \geq x_i^2 x_j > x_j^2 > 0 x_i^2
\]

so that \( x_i \) would be accepted
as a \( 1 \times 1 \) pivot and no pivoting
would be done. This property would hold
throughout the factorization which
implies that the symmetric indefinite
factorization reduces to

\[
A = LDL^T \quad \text{with } L \text{ unit lower}
\]
\[
\text{triangular and } D
\]
\[
\text{diagonal (pos. def)}
\]

without any pivoting.

It can be shown that the pivot
selection strategy assures a
boundedness of ideal growth in
the update
\[
a_{ij} \rightarrow a_{ij} - \frac{a_{ik} a_{kj}}{x_k}
\]
at each step, giving a bound on Growth Factor.