A direct orthogonal sparse static methodology for a finite continuation hybrid LP solver

Pablo Guerrero-García and Ángel Santos-Palomo

Dpt. Applied Mathematics, University of Málaga

June 22nd, 2004 (Technical Report MA-04/03)

Plan of the talk

✓**Part I: Proposed LP framework**

0. Notation
1. Finite continuation Phase-I and crossover
2. Primal NSA for Phase-II

Part II: Sparse orthogonal factorizations

3. Original technique
4. The modification used in Phase-I
5. Additional features

Appendix: Updating and downdating

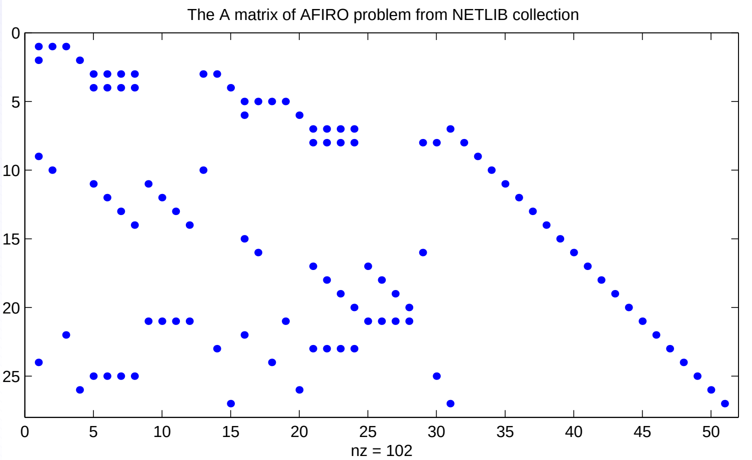
6. Original technique: Updating
7. Original technique: Updating example
8. Original technique: Downdating
9. Original technique: Downdating example

0. Notation

Unsymmetric primal-dual LP pair (non-standard notation):

$$(P) \quad \min \ \ell(x) \doteq c^T x, \quad x \in \mathbb{R}^n \qquad \text{s.t.} \quad A^T x \geq b$$

$$(D) \quad \max \ \mathcal{L}(y) \doteq b^T y, \quad y \in \mathbb{R}^m \qquad \text{s.t.} \quad Ay = c, \quad y \geq O$$


The A matrix of AFIRO problem from NETLIB collection

- $A \in \mathbb{R}^{n\times m}$, $m \geq n$, $\operatorname{rank}(A) = n$, $\operatorname{rank}(A_k) = m_k$ (Phase-II only)
- $m_k \doteq |\mathcal{B}^{(k)}|$, $m_k \leq n$ (Phase-II only), $\mathcal{N}^{(k)} \doteq [1:m]\backslash\mathcal{B}^{(k)}$
- $A_k \doteq A(:,\mathcal{B}^{(k)}) \in \mathbb{R}^{n\times m_k}$, $N_k \doteq A(:,\mathcal{N}^{(k)}) \in \mathbb{R}^{n\times(m-m_k)}$
- $a_i \doteq A(:,i)$, $a_i^T \doteq A^T(i,:)$
- $Z_k \in \mathbb{R}^{n\times(n-m_k)}$ basis (not orthonormal) of $\mathcal{N}(A_k^T)$
- $\mathcal{F} \doteq \{x \in \mathbb{R}^n : A^T x \geq b\}$, $\mathcal{G} \doteq \{y \in \mathbb{R}^m : Ay = c, y \geq O\}$

1. Finite continuation Phase-I and crossover (I)

- Obtain $y \in \mathcal{G}$ by minimizing a piecewise quadratic (Nielsen, 1998)
- Have to solve sequence of sparse linear systems with coefficient matrix $\sigma I_n + A_k A_k^T$ (constant $\sigma > 0$), as shown by (Nielsen, 1999, pp. 97–98)
- Crossover to obtain a vertex of $\mathcal{G}$ for Phase-II to start with
- Ascent direction projecting $b_\mathcal{B}$ onto $\mathcal{N}(A_k)$, deletion or exchange then performed, and update of $y \in \mathcal{G}$ after usual primal min-ratio
- Final (full-column rank) A_k after crossover not restricted to be square:

Example 1 (Murty, 83, p. 122): $y^{(0)} = [1; 2; 3; 4; 0; 0; 0]$,

$$\left[\begin{array}{c|c} b^T & \\ \hline A & c \end{array}\right] = \left[\begin{array}{rrrrrrr|r} 10 & -4 & -6 & -2 & -4 & -8 & -10 & \\ \hline 1 & 1 & 3 & 4 & 0 & 0 & 0 & 28 \\ 0 & -1 & -1 & -2 & 1 & 0 & 0 & -13 \\ 0 & 1 & 1 & 2 & 0 & 1 & 0 & 13 \\ 1 & 0 & 2 & 2 & 0 & 0 & -1 & 15 \end{array}\right],$$

k	$\mathcal{B}^{(k)}$	$\mathcal{L}^{(k)}$
0	$\{1, 2, 3, 4\}$	-24.0
1	$\{1, 2, 4\}$	$+10.9$
2	$\{1, 2\}$	$+98.0$

1. Finite continuation Phase-I and crossover (and II)

Example 2 (Murty, 88, pp. 475–476): $y^{(0)} = [5;12;13;1;2;0;0]/2$,

$$\left[\begin{array}{c|c} b^T & \\ \hline A & c \end{array}\right] = \left[\begin{array}{ccccccc|c} 10 & -4 & -6 & -2 & -4 & -8 & -10 & \\ \hline 1 & 0 & 0 & 1 & 0 & 1 & -1 & 3 \\ 0 & 1 & 0 & 0 & -1 & 2 & -1 & 5 \\ 0 & 0 & 1 & -1 & 1 & 1 & -2 & 7 \end{array}\right],$$

k	$\mathcal{B}^{(k)}$	$\mathcal{L}^{(k)}$
0	$\{1,2,3,4,5\}$	-43.0
1	$\{1,2,3,5\}$	-33.3
2	$\{1,2,3\}$	-32.0

Example 3 (Murty, 83, p. 123, Ex. 3.24a): $y^{(0)} = [1;2;3;4;5;0;0]$,

$$\left[\begin{array}{c|c} b^T & \\ \hline A & c \end{array}\right] = \left[\begin{array}{ccccccc|c} -1 & -1 & -1 & -2 & -2 & -2 & -4 & \\ \hline 1 & 1 & 0 & 2 & 1 & 0 & 1 & 16 \\ 0 & 1 & 1 & 1 & 2 & 0 & 5 & 19 \\ 1 & 0 & 1 & 1 & 1 & 2 & 2 & 13 \end{array}\right],$$

k	$\mathcal{B}^{(k)}$	$\mathcal{L}^{(k)}$
0	$\{1,2,3,4,5\}$	-24.0

Example 4 (Murty, 83, p. 123, Ex. 3.24b): $y^{(0)} = [1;2;3;4;5;0;0]$,

$$\left[\begin{array}{c|c} b^T & \\ \hline A & c \end{array}\right] = \left[\begin{array}{ccccccc|c} -1 & -1 & -1 & -2 & -2 & 0 & -4 & \\ \hline 1 & 1 & 0 & 2 & 1 & 0 & 1 & 16 \\ 0 & 1 & 1 & 1 & 2 & 0 & 5 & 19 \\ 1 & 0 & 1 & 1 & 1 & 2 & 2 & 13 \end{array}\right],$$

k	$\mathcal{B}^{(k)}$	$\mathcal{L}^{(k)}$
0	$\{1,2,3,4,5\}$	-24.0
1	$\{2,3,4,5,6\}$	-23.2
2	$\{2,4,5,6\}$	-22.4
3	$\{2,5,6\}$	-19.0

2. Primal NSA for Phase-II

S0. *Initialize.*
Let $k \leftarrow 0$ and choose $y^{(0)} \in \mathcal{G}$ with basic set $\mathcal{B}^{(0)}$ and basis A_0.

S1. *Check for optimality.*
Compute solution $x^{(k)}$ of $A_k^T x = b_{\mathcal{B}}$ and residues $r_{\mathcal{N}}^{(k)} = N_k^T x^{(k)} - b_{\mathcal{N}}$ of non-basic constraints ($r_{\mathcal{B}}^{(k)} = O$). If $r_{\mathcal{N}}^{(k)} \geq O$ stop, $x^{(k)}$ optimal.

S2. *Pick non-basic to add and define (if possible) search dir for basics.*
Choose $j \in 1\colon (m - m_k)$ such that $r_{\mathcal{N}_j}^{(k)} < 0$. With $p \doteq \mathcal{N}_j^{(k)}$, determine whether search dir. $d_{\mathcal{B}}^{(k)} \in \mathbb{R}^{m_k}$ for basics exists ($d_{\mathcal{N}}^{(k)} = -e_j \in \mathbb{R}^{m-m_k}$) by checking compatibility of $A_k d_{\mathcal{B}} = a_p$:

(a) Compat. sist. if $Z_k^T a_p = O$: obtain $d_{\mathcal{B}}^{(k)}$ and go on to **S3**.

(b) Incompat. sist. if $Z_k^T a_p \neq O$: $a_p \notin \mathcal{R}(A_k)$ can be added to $\mathcal{B}^{(k)}$ without deleting constraints from $\mathcal{B}^{(k)}$; indicated by $i \leftarrow 0$, $q \leftarrow \varnothing$ and going on to **S5** with $\tau = 0$.

2. Primal NSA for Phase-II (cont.)

S3. *Calculate maximum feasible step along search dir.*
For $i \in 1\colon m_k$ calculate step τ_i such that ith basic var would be zero:

$$\tau_i = \begin{cases} y_{\mathcal{B}_i}^{(k)} / d_{\mathcal{B}_i}^{(k)} & \text{, if } d_{\mathcal{B}_i}^{(k)} > 0 \\ +\infty & \text{, if } d_{\mathcal{B}_i}^{(k)} \leq 0 \end{cases} \quad ; \quad \tau = \min_{1 \leq i \leq m_k} \tau_i$$

S4. *Test for unbounded solution and pick basic to delete.* If $\tau = +\infty$, $\mathcal{L}$ unbounded above in $\mathcal{G}$ ($\mathcal{F} = \varnothing$); else pick $i \in 1\colon m_k$ with $\tau = \tau_i$.

S5. *Prepare next iteration.* If $i \neq 0$ then

$$q \doteq \mathcal{B}_i^{(k)} \quad ; \quad y_{\mathcal{B}_i}^{(k)} \leftarrow \varnothing \quad ; \quad d_{\mathcal{B}_i}^{(k)} \leftarrow \varnothing \quad ; \quad \mathcal{B}_i^{(k)} \leftarrow \varnothing$$

and delete ith column from A_k. Then let $\mathcal{N}_j^{(k)} \leftarrow \varnothing$,

$$y_{\mathcal{B}}^{(k)} \leftarrow \begin{bmatrix} y_{\mathcal{B}}^{(k)} - \tau d_{\mathcal{B}}^{(k)} \\ \tau \end{bmatrix} \quad ; \quad \mathcal{B}^{(k)} \leftarrow [\mathcal{B}^{(k)},\ p] \quad ; \quad \mathcal{N}^{(k)} \leftarrow [\mathcal{N}^{(k)},\ q].$$

Append a_p to A_k and let $k \leftarrow k+1$. Go back to **S1**.

Plan of the talk

Part I: Proposed LP framework

0. Notation
1. Finite continuation Phase-I and crossover
2. Primal NSA for Phase-II

✓**Part II: Sparse orthogonal factorizations**

3. Original technique
4. The modification used in Phase-I
5. Additional features

Appendix: Updating and downdating

6. Original technique: Updating
7. Original technique: Updating example
8. Original technique: Downdating
9. Original technique: Downdating example

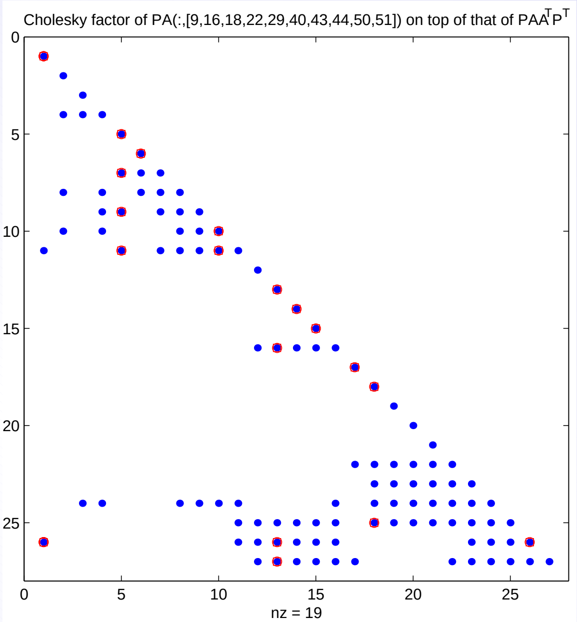
3. Original technique (I)

- Orthogonal factoriz. used without explicit orthogonal (dense) factor, work on top of a sparse triangular matrix (Cholesky factor R of AA^T)

- Phase-II: $\mathrm{rank}(A_k) = m_k \le n$

- Crossover use same data structure

- Sparse QR factoriz. of A_k^T (± rows). Column order of A_k has no impact on density of Cholesky factor R_k of $A_kA_k^T$, (Saunders, 72) techniques for square matrices adapted to matrices with more rows than columns, using static data structure of (George and Heath, 80) but allowing LINPACK-like row downdating on it



- If A_k column subset of fixed $A \in \mathbb{R}^{n\times m}$ with $m \ge n$, then sparsity structure of $A_kA_k^T$ subset of that of AA^T (George and Heath, 80), “a priori” permutation P of rows of A to improve sparsity of R

3. Original technique (II)

$R_k\Pi_k$ row-echelon, perm. upper trapez. with $\Pi_k^T = [e_1, e_3, e_6, e_2, e_4, e_5, e_7]$

$$\mathtt{qr}\underbrace{\left(\begin{bmatrix} \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \end{bmatrix}\right)}_{A_k^T} = \underbrace{\begin{bmatrix} \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{X} & \mathtt{X} \end{bmatrix}}_{Q_k} \cdot \underbrace{\begin{bmatrix} \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{0} & \mathtt{0} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{X} & \mathtt{X} \end{bmatrix}}_{R_k\cdot\Pi_k}$$

Let sparse $A_k \in \mathbb{R}^{n\times m_k}$ with $\mathrm{rank}(A_k) = m_k \leq n$, there exists permutation Π_k^T of A_k^T columns *implicitly defined* by staircase shape of structure,

$$A_k^T\Pi_k^T = Q_k R_k \doteq Q_k \left[\, R_{\mathrm{i}} \quad R_{\mathrm{d}} \,\right]$$

$$(Q_k, R_{\mathrm{i}} \in \mathbb{R}^{m_k\times m_k}) \qquad \text{and} \qquad (R_{\mathrm{d}} \in \mathbb{R}^{m_k\times(n-m_k)})$$

with R_k upper trapez., R_{i} upper triang. and regular, Q_k orthogonal:

$$\Pi_k A_k A_k^T \Pi_k^T = R_k^T \underbrace{Q_k^T Q_k}_{I} R_k = R_k^T R_k$$

Updating and downdating ($\mathcal{O}(n^2)$) this Cholesky factorization can be done (Santos and Guerrero, 02, submitted to IMA J. Numerical Analysis)

3. Original technique (and III)

- Columns of R_k^T base of $\mathcal{R}(\Pi_k A_k)$ and Z base of $\mathcal{N}(A_k^T)$:

$$Z = \Pi_k^T \begin{bmatrix} Z_1^T \\ I \end{bmatrix}, \qquad \text{where} \qquad Z_1 \doteq -R_{\mathrm{d}}^T R_{\mathrm{i}}^{-T}$$

$$v \in \mathcal{R}(A_k) \quad \Leftrightarrow \quad \forall s \in \mathcal{N}(A_k^T)\ ,\ v^T s = 0 \quad \Leftrightarrow \quad Z^T v = \mathrm{O}$$

$$v \notin \mathcal{R}(A_k) \quad \Leftrightarrow \quad Z^T v \neq \mathrm{O} \qquad (\|Z^T v\| < \epsilon)$$

$$\Pi_k v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},\ v_1 \in \mathbb{R}^{m_k} \quad \Rightarrow \quad Z^T v = Z_1 v_1 + v_2 = -R_{\mathrm{d}}^T R_{\mathrm{i}}^{-T} v_1 + v_2$$

- Basic solution of $A_k^T x = b_{\mathcal{B}}$ by special "seminormal equations":

$$A_k A_k^T x = \Pi_k^T R_k^T R_k \Pi_k x = A_k b_{\mathcal{B}} \qquad \Rightarrow \qquad R_k^T (R_k(\Pi_k x)) = \Pi_k (A_k b_{\mathcal{B}}),$$

and analogously for $A_k y_{\mathcal{B}} = c$ with $y_{\mathcal{B}} \doteq A_k^T z$ but with $A_k A_k^T z = c$

$$A_k A_k^T z = \Pi_k^T R_k^T R_k \Pi_k z = c \qquad \Rightarrow \qquad R_k^T (R_k(\Pi_k z)) = \Pi_k c.$$

- Zero diagonals of R temporarily set to 1 before triangsolve (Heath, 82)
- Compatibility of $A_k^T x = b_{\mathcal{B}}$ checked at each crossover iteration

4. The modification used in Phase-I

We have to work in Phase-I with the Cholesky factor R_k of $\sigma I_n + A_k A_k^T > 0$ (constant σ), instead of with the Cholesky factor R_k of $A_k A_k^T \geq 0$:

Theorem 1 *Cholesky factor of* $\sigma I_n + A_k A_k^T$ *coincides (barring signs) with triangular factor of QR factoriz. of* $\left[\, A_k, \sqrt{\sigma} I_n \,\right]^T$, *namely*

$$\sigma I_n + A_k A_k^T = R_k^T R_k \qquad \Leftrightarrow \qquad \begin{bmatrix} A_k^T \\ \sqrt{\sigma} I_n \end{bmatrix} = Q_k \begin{bmatrix} R_k \\ O_k \end{bmatrix}.$$

Proof

$$R_k^T R_k = \left[\, R_k^T, O_k \,\right] Q_k^T Q_k \begin{bmatrix} R_k \\ O_k \end{bmatrix} = \left[\, A_k, \sqrt{\sigma} I_n \,\right] \begin{bmatrix} A_k^T \\ \sqrt{\sigma} I_n \end{bmatrix} = \sigma I_n + A_k A_k^T$$

- Full column rank of A_k requisite in original technique can be relaxed
- Cholesky factor structure of AA^T coincides with that of $\sigma I_n + AA^T$
- Initialization from $R = \sqrt{\sigma} I_n$ in diagonal of static structure
- Each row of A_k^T linearly dependent of rows already present in structure

5. Additional features

- Original technique MATLAB toolbox devised by (Guerrero, 02, PhD)
- Availability of SPARSPAK++ package (George and Liu, 99, SIMAX) would imply rapid prototyping of low-level C++ implementation
- Parallelizability and decrease of computational effort when refactorization has to be done (currently with sparse row-sequential option of MATLAB `qr` command, no source code available, row order of A_k^T)
- Recomputing factorization from scratch every iteration would prohibitively imply $\mathcal{O}(n^3)$ per iteration
- Same sparse static data structure used in all phases!!

Plan of the talk

Part I: Proposed LP framework

0. Notation
1. Finite continuation Phase-I and crossover
2. Primal NSA for Phase-II

Part II: Sparse orthogonal factorizations

3. Original technique
4. The modification used in Phase-I
5. Additional features

✓**Appendix: Updating and downdating**

6. Original technique: Updating
7. Original technique: Updating example
8. Original technique: Downdating
9. Original technique: Downdating example

6. Original technique: Updating (I)

Our problem can be stated as

$$A_{k+1}^T = \left[\begin{array}{c} A_k^T \\ a_p^T \end{array} \right] \quad \text{and} \quad A_k^T \Pi_k^T = Q_k R_k \quad \Rightarrow \quad ¿A_{k+1}^T \Pi_{k+1}^T = Q_{k+1} R_{k+1}?$$

Since $A_k^T = Q_k R_k \Pi_k$, we can write

$$A_{k+1}^T = \left[\begin{array}{c} A_k^T \\ a_p^T \end{array} \right] = \left[\begin{array}{cc} Q_k & \mathrm{O} \\ \mathrm{O}^T & 1 \end{array} \right] \left[\begin{array}{c} R_k \Pi_k \\ a_p^T \end{array} \right],$$

so we have

$$\left[\begin{array}{cc} Q_k^T & \mathrm{O} \\ \mathrm{O}^T & 1 \end{array} \right] \left[\begin{array}{c} A_k^T \\ a_p^T \end{array} \right] = \left[\begin{array}{c} R_k \Pi_k \\ a_p^T \end{array} \right] = \left[\begin{array}{ccccccc} \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{0} & \mathtt{0} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \end{array} \right].$$

6. Original technique: Updating (and II)

Let f_j be the *index* of the column of Π_k^T where e_j appears, e.g.,

$$f_1 = 1, \quad f_3 = 2, \quad f_6 = 3, \quad \text{and} \quad f_2 = 4 > 3 \doteq m_k.$$

For each $j \in 1\colon n$, annihilate jth element (if nonzero) of updated version of a_p^T by a Givens rotation $G(f_j, m_k{+}1)^T \in \mathbb{R}^{(m_k+1)\times(m_k+1)}$ only if $f_j \le m_k$, since otherwise we are done. Hence, after m_k rotations (worse case)

$$G^T \doteq G(f_n, m_k + 1)^T \cdots G(f_2, m_k + 1)^T G(f_1, m_k + 1)^T,$$

$$F^T G^T \begin{bmatrix} R_k \Pi_k \\ a_p^T \end{bmatrix} \doteq R_{k+1} \Pi_{k+1},$$

$$A_{k+1}^T \Pi_{k+1}^T = \begin{bmatrix} Q_k & \mathrm{O} \\ \mathrm{O}^T & 1 \end{bmatrix} G F F^T G^T \begin{bmatrix} R_k \Pi_k \\ a_p^T \end{bmatrix} \Pi_{k+1}^T \doteq Q_{k+1} R_{k+1}.$$

R_k updated even if Q_k not available; both Π_{k+1} and F^T implicitly defined.

7. Original technique: Updating example (I)

$$A^T = \begin{bmatrix} \times & & & & & \times \\ & \times & & \times & & \\ 1 & & -1 & & & \\ & \times & & \times & & \times \\ & & & & 1 & 1 \\ 1 & & & & & 2 \\ & & -1 & -1 & & \\ & \times & & \times & & \end{bmatrix}, \qquad R^T = \begin{bmatrix} \times & & & & & \\ & \times & & & & \\ \times & & \times & & & \\ & \times & \times & \times & & \\ & & & & \times & \\ \times & \times & \times & \times & \times & \times \end{bmatrix}.$$

Rows a_6^T and a_7^T are rotated into R, and now a_3^T is going to be rotated:

$$[R^T|a_3] = \left[\begin{array}{cccccc|c} 1 & & & & & & 1 \\ & \times & & & & & \\ \times & & -1 & & & & -1 \\ & \times & -1 & \times & & & \\ & & & & \times & & \\ 2 & \times & \times & \times & \times & \times & \end{array}\right] \text{ and } [\Pi_2^T R_2^T|a_3] = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 0 & 0 \\ \mathbf{0} & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & \mathbf{0} & 0 \end{array}\right].$$

7. Original technique: Updating example (and II)

Postmultiplying by $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and by $\begin{bmatrix} \sqrt{6}/3 & -1/\sqrt{3} \\ 1/\sqrt{3} & \sqrt{6}/3 \end{bmatrix}$, first columns 1 and 3 and then columns 2 and 3:

$$\left[\begin{array}{cc|c} \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ -1/\sqrt{2} & -1 & -1/\sqrt{2} \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ \sqrt{2} & \mathbf{0} & -\sqrt{2} \end{array}\right] \quad \text{and} \quad \Pi_3^T R_3^T = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ -1/\sqrt{2} & -\sqrt{6}/2 & 0 \\ 0 & -\sqrt{6}/3 & 1/\sqrt{3} \\ 0 & 0 & 0 \\ \sqrt{2} & -\sqrt{6}/3 & -2/\sqrt{3} \end{bmatrix}.$$

Finally, addition of a_5^T entails no rotations:

$$\Pi_4^T R_4^T = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1/\sqrt{2} & -\sqrt{6}/2 & 0 & 0 \\ 0 & -\sqrt{6}/3 & 1/\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \\ \sqrt{2} & -\sqrt{6}/3 & -2/\sqrt{3} & 1 \end{bmatrix}.$$

8. Original technique: Downdating (I)

Our problem can be stated as

$$A_k^T = \left[\begin{array}{c} a_q^T \\ A_{k+1}^T \end{array} \right] \quad \text{and} \quad A_k^T \Pi_k^T = Q_k R_k \quad \Rightarrow \quad ¿A_{k+1}^T \Pi_{k+1}^T = Q_{k+1} R_{k+1}?$$

This fits well into the Linpack ($m_k > n$) Cholesky downdating algorithm since $\Pi_k^T R_k^T q = a_q$ *is also compatible with only one solution.* Apply to

$$\bar{R}_k \doteq \left[\begin{array}{cc} \delta_n & \mathrm{O}^T \\ q & R_k \Pi_k \end{array} \right] = \left[\begin{array}{cccccccc} \mathtt{X} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{0} & \mathtt{0} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} & \mathtt{X} \\ \mathtt{X} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{X} & \mathtt{X} \end{array} \right]$$

several Givens rotations $G(1,j)^T \in \mathbb{R}^{(m_k+1)\times(m_k+1)}$ with $j \in m_k{+}1 : -1 : 2$

$$F^T G^T \bar{R}_k \doteq F^T G(1,2)^T \cdots G(1,m_k)^T G(1,m_k+1)^T \bar{R}_k \doteq \left[\begin{array}{c} \bar{R}_{k+1} \\ \mathrm{O}^T \end{array} \right],$$

$$\bar{R}_{k+1} \doteq \left[\begin{array}{cc} \alpha & \alpha v^T \Pi_{k+1} \\ \mathrm{O} & R_{k+1} \Pi_{k+1} \end{array} \right] \qquad (\alpha = \pm 1).$$

8. Original technique: Downdating (and II)

Since

$$G \text{ and } F \text{ orthogonal} \qquad \Rightarrow \qquad \bar{R}_{k+1}^T \bar{R}_{k+1} = \bar{R}_k^T \bar{R}_k,$$

$$\begin{bmatrix} \alpha & \mathrm{O}^T \\ \alpha \Pi_{k+1}^T v & \Pi_{k+1}^T R_{k+1}^T \end{bmatrix} \begin{bmatrix} \alpha & \alpha v^T \Pi_{k+1} \\ \mathrm{O} & R_{k+1} \Pi_{k+1} \end{bmatrix} = \begin{bmatrix} \delta_n & q^T \\ \mathrm{O} & \Pi_k^T R_k^T \end{bmatrix} \begin{bmatrix} \delta_n & \mathrm{O}^T \\ q & R_k \Pi_k \end{bmatrix}.$$

Multiplying out and comparing blocks

$$\begin{bmatrix} 1 & v^T \Pi_{k+1} \\ \Pi_{k+1}^T v & \Pi_{k+1}^T (R_{k+1}^T R_{k+1} + v v^T) \Pi_{k+1} \end{bmatrix} = \begin{bmatrix} \delta_n^2 + \|q\|_2^2 & q^T R_k \Pi_k \\ \Pi_k^T R_k^T q & \Pi_k^T R_k^T R_k \Pi_k \end{bmatrix},$$

we have that

$$\delta_n^2 = 1 - \|q\|_2^2 = 0 \qquad \text{and} \qquad \Pi_{k+1}^T v = \Pi_k^T R_k^T q = a_q,$$

thus

$$A_{k+1} A_{k+1}^T = \Pi_{k+1}^T R_{k+1}^T R_{k+1} \Pi_{k+1} = \Pi_k^T R_k^T R_k \Pi_k - a_q a_q^T = A_k A_k^T - a_q a_q^T.$$

9. Original technique: Downdating example (I)

Now we want to delete a_7^T. First solve the compatible system

$$\Pi_4^T R_4^T q = a_7 \doteq [0; 0; -1; -1; 0; 0]$$

to obtain $q = [0; \sqrt{6}/3; -1/\sqrt{3}; 0]$ (overlaying $\Pi_4^T R_4^T$ on I_6 to obtain $[0; 0; \sqrt{6}/3; -1/\sqrt{3}; 0; 0]$ and select 1st, 3rd, 4th and 5th components):

$$\bar{R}_4^T \doteq \begin{bmatrix} \delta_n & q^T \\ \mathrm{O} & \Pi_4^T R_4^T \end{bmatrix} = \left[\begin{array}{c|cccc} 0 & 0 & \sqrt{6}/3 & -1/\sqrt{3} & 0 \\ \hline 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1/\sqrt{2} & -\sqrt{6}/2 & 0 & 0 \\ 0 & 0 & -\sqrt{6}/3 & 1/\sqrt{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & \sqrt{2} & -\sqrt{6}/3 & -2/\sqrt{3} & 1 \end{array}\right].$$

9. *Original technique: Downdating example (and II)*

Postmultiplying by $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and by $\begin{bmatrix} -1/\sqrt{3} & \sqrt{6}/3 \\ -\sqrt{6}/3 & -1/\sqrt{3} \end{bmatrix}$, first columns 1 and 4 and then columns 1 and 3,

$$\left[\begin{array}{c|cccc} 1/\sqrt{3} & 0 & \sqrt{6}/3 & 0 & 0 \\ \hline 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1/\sqrt{2} & -\sqrt{6}/2 & 0 & 0 \\ -1/\sqrt{3} & 0 & -\sqrt{6}/3 & \mathbf{0} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 2/\sqrt{3} & \sqrt{2} & -\sqrt{6}/3 & \mathbf{0} & 1 \end{array}\right], [\bar{R}_5^T\ \mathrm{O}]F \doteq \left[\begin{array}{c|cccc} -1 & 0 & 0 & 0 & 0 \\ \hline 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 1 & 0 & \mathbf{0} & \mathbf{0} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & \sqrt{2} & \sqrt{2} & \mathbf{0} & 1 \end{array}\right].$$

Zero column just created has not appeared in the last position, hence $F \neq I$ in general; static structure avoids explicit computation of F.