Automation in Computational Biology

Diego Fabregat and **Paolo Bientinesi**
In collaboration with Prof. Yurii Aulchenko

AICES, RWTH Aachen
pauldj@aices.rwth-aachen.de

9th International Conference on
Parallel Processing and Applied Mathematics
September 11-14, 2011
Torun, Poland

The problem - biology

Mixed models: $b = \left(X^T M^{-1} X\right)^{-1} X^T M^{-1} y$

Genome-wide association analysis

- y: phenotype
  (outcome; vector of observations)
  E.g.: height, blood pressure for a set of people
- X: genome measurements and covariates
  (design matrix; predictors)
  E.g.: sex and age over height
- M: dependencies between observations
  E.g.: tall parents have tall children
- b: relation between a variation in the outcome (y) and a variation in the genome sequence (X)

The problem - linear algebra

$$b = \left(X^T M^{-1} X\right)^{-1} X^T M^{-1} y$$

Linear regression with non-independent outcomes

Inputs

- $M \in \mathcal{R}^{n \times n}, \quad SPD(M), \quad n \in [10^3, \ldots, 10^4]$
- $X \in \mathcal{R}^{n \times p}, \quad p \in [1, \ldots, 20]$
- $y \in \mathcal{R}^n$

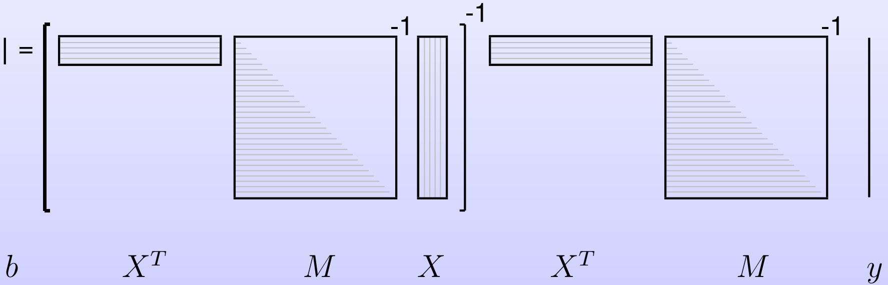
Output

- $b \in \mathcal{R}^p$

$\star$To be repeated thousands of times$\star$

Mixed Models

Shape visualization



Note

What this is talk is **NOT** about

$$\{LU = A,\ A, L, U \in \mathbb{R}^{n\times n},\ \texttt{LowerTriUni}(L),\ \texttt{UpperTri}(U)\} \Rightarrow$$

Partition $A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$

where A_{TL} is 0×0

While $size(A_{TL}) < size(A)$ **do**

Repartition

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

$A_{01} := L_{00}^{-1} A_{01}$

$A_{10} := A_{10} U_{00}^{-1}$

$A_{11} := \mathsf{LU}(A_{11} - A_{10} A_{01})$

Continue

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

endwhile

Note

What this is talk is **NOT** about

$$\{LU = A,\ A, L, U \in \mathbb{R}^{n\times n},\ \texttt{LowerTriUni}(L),\ \texttt{UpperTri}(U)\} \Rightarrow$$

Partition $A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$

where A_{TL} is 0×0

While $size(A_{TL}) < size(A)$ **do**

Repartition

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

$A_{01} := L_{00}^{-1} A_{01}$

$A_{10} := A_{10} U_{00}^{-1}$

$A_{11} :=$ LU$(A_{11} - A_{10} A_{01})$

Continue

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

endwhile

$b = \left(X^T M^{-1} X\right)^{-1} X^T M^{-1} y$

Matlab: `inv( X' * inv( M ) * X ) * X' * inv( M ) * y`

How to compute b?

- High-performance
- Multithreading

$b = \left(X^T M^{-1} X\right)^{-1} X^T M^{-1} y$

Matlab: `inv( X' * inv( M ) * X ) * X' * inv( M ) * y`

How to compute b?

- High-performance
- Multithreading
- Available: BLAS, LAPACK
  - Definitely not BLAS
  - No direct call to LAPACK,
    but closely related routines are available
    E.g., `DGELS` computes $\left(X^T X\right)^{-1} X^T y$
- Objective:
  Expressing b in terms of BLAS and LAPACK calls

Example: a couple of algorithms

$$b = \left(X^T\ M^{-1}\ X\right)^{-1} X^T\ M^{-1}\ y$$

Algorithm 1

Example: a couple of algorithms

$$b = \left(X^T \, \mathbf{M^{-1}} \, X\right)^{-1} X^T \, \mathbf{M^{-1}} \, y$$

Algorithm 1

1. $M \leftarrow M^{-1}$

Example: a couple of algorithms

$$b = \left(\mathbf{X^T}\,\mathbf{M}\,X\right)^{-1}\mathbf{X^T}\,\mathbf{M}\,y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\ M$

Example: a couple of algorithms

$$b = (\mathbf{aux_1}\ \mathbf{X})^{-1}\, aux_1\ y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\ M$
3. $aux_2 \leftarrow aux_1\ X$

Example: a couple of algorithms

$$b = (\mathbf{aux_2})^{-\mathbf{1}} aux_1 \; y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T \; M$
3. $aux_2 \leftarrow aux_1 \; X$
4. $aux_2 \leftarrow aux_2^{-1}$

Example: a couple of algorithms

$$b = aux_2 \, \mathbf{aux_1} \, \mathbf{y}$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T \, M$
3. $aux_2 \leftarrow aux_1 \, X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1 \, y$

Example: a couple of algorithms

$$b = \mathbf{aux_2\ y}$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\ M$
3. $aux_2 \leftarrow aux_1\ X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1\ y$
6. $b \leftarrow aux_2\ b$

Example: a couple of algorithms

$$b = \left(X^T \, M^{-1} \, X\right)^{-1} X^T \, M^{-1} \, y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T \, M$
3. $aux_2 \leftarrow aux_1 \, X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1 \, y$
6. $b \leftarrow aux_2 \, b$

Example: a couple of algorithms

$$b = \left(X^T \, \mathbf{M}^{-1} \, X\right)^{-1} X^T \, \mathbf{M}^{-1} \, y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T \; M$
3. $aux_2 \leftarrow aux_1 \; X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1 \; y$
6. $b \leftarrow aux_2 \; b$

Algorithm 2

1. $LL^T = M$

Example: a couple of algorithms

$$b = \left(X^T \left(\mathbf{L}\mathbf{L}^\mathbf{T}\right)^{-1} X\right)^{-1} X^T \left(\mathbf{L}\mathbf{L}^\mathbf{T}\right)^{-1} y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\ M$
3. $aux_2 \leftarrow aux_1\ X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1\ y$
6. $b \leftarrow aux_2\ b$

Algorithm 2

1. $LL^T = M$

Example: a couple of algorithms

$$b = \left(X^T\, L^{-T}\, L^{-1}\, X\right)^{-1} X^T\, L^{-T}\, L^{-1}\, y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\, M$
3. $aux_2 \leftarrow aux_1\, X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1\, y$
6. $b \leftarrow aux_2\, b$

Algorithm 2

1. $LL^T = M$

Example: a couple of algorithms

$$b = \left(X^T\, L^{-T}\, \mathbf{L^{-1}}\, \mathbf{X}\right)^{-1} X^T\, L^{-T}\, L^{-1}\, y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\, M$
3. $aux_2 \leftarrow aux_1\, X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1\, y$
6. $b \leftarrow aux_2\, b$

Algorithm 2

1. $LL^T = M$
2. $X \leftarrow L^{-1} X$

Example: a couple of algorithms

$$b = \left(X^T\ X\right)^{-1} X^T\ \mathbf{L^{-1}}\ \mathbf{y}$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\ M$
3. $aux_2 \leftarrow aux_1\ X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1\ y$
6. $b \leftarrow aux_2\ b$

Algorithm 2

1. $LL^T = M$
2. $X \leftarrow L^{-1} X$
3. $y \leftarrow L^{-1} y$

Example: a couple of algorithms

$$b = \left(X^T\ X\right)^{-1} X^T\ y$$

Algorithm 1

1. $M \leftarrow M^{-1}$
2. $aux_1 \leftarrow X^T\ M$
3. $aux_2 \leftarrow aux_1\ X$
4. $aux_2 \leftarrow aux_2^{-1}$
5. $b \leftarrow aux_1\ y$
6. $b \leftarrow aux_2\ b$

Algorithm 2

1. $LL^T = M$
2. $X \leftarrow L^{-1}X$
3. $y \leftarrow L^{-1}y$
4. $b = \texttt{DGELS}(X, y)$

Automation

CL1CK (ver. 1.0)

- Symbolic system for linear algebra equations
- Written in Mathematica (D. Fabregat)

Automation

Cl1ck (ver. 1.0)

- Symbolic system for linear algebra equations
- Written in Mathematica (D. Fabregat)
- **Pattern matching** & rewrite rules
- Constructive: no exhaustive search
- Knowledge: 1) encoded + 2) derived

Automation

CL1CK (ver. 1.0)

- Symbolic system for linear algebra equations
- Written in Mathematica (D. Fabregat)
- **Pattern matching** & rewrite rules
- Constructive: no exhaustive search
- Knowledge: 1) encoded + 2) derived
- Input: equation and operands' description
- Output: algorithms and code
- (Not a theorem prover)

CL1CK: input

```
precond = {
  {X, {“Input”, “Matrix”, “FullRank”} }
  {y, {“Input”, “Vector” } }
  {M, {“Input”, “Matrix”, “SPD”, “Lower” } }
  {b, {“Output”, “Vector” } }
};
sizeAssump = { row[X] > col[X] };
```

```
postcond = {
  equal[b,
    times[
      inv[ times[ trans[X], inv[M], X ] ],
      trans[X],
      inv[M],
      y ]
  ]
};
```

CL1CK: output algorithms (selected)

$$b = \left(X^T\ M^{-1}\ X\right)^{-1} X^T\ M^{-1}\ y$$

Algorithm 1

$L1\ L1^T == m$

$tmp3 == x^T\ L1^{-T}$

$tmp8 == L1^{-1}\ y$

$tmp14 == inv\left[tmp3\ tmp3^T\right]\ tmp3\ tmp8$

$b == tmp14$

CL1CK: output algorithms (selected)

$$b = \left(X^T\ M^{-1}\ X\right)^{-1} X^T\ M^{-1}\ y$$

Algorithm 1

$L1\ L1^T == m$

$tmp3 == x^T\ L1^{-T}$

$tmp8 == L1^{-1}\ y$

$tmp14 == inv\left[tmp3\ tmp3^T\right]\ tmp3\ tmp8$

$b == tmp14$

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $y \leftarrow L^{-1} y$
4. $b = \texttt{DGELS}(X, y)$

CL1CK: output algorithms (selected)

$$b = \left(X^T\ M^{-1}\ X\right)^{-1} X^T\ M^{-1}\ y$$

Algorithm 2

```
L1 L1^T == m
tmp3 == x^T L1^-T
Q8 R8 == tmp3^T
tmp10 == L1^-1 y
tmp11 == Q8^T tmp10
tmp16 == R8^-1 tmp11
b == tmp16
```

CL1CK: output algorithms (selected)

$$b = \left(X^T\, \mathbf{M}^{-1}\, X\right)^{-1} X^T\, \mathbf{M}^{-1}\, y$$

Algorithm 2

$\mathtt{L1}\,\mathtt{L1}^{\mathtt{T}} := \mathtt{m}$

$\mathtt{tmp3} := \mathtt{x}^{\mathtt{T}}\,\mathtt{L1}^{\mathtt{-T}}$

$\mathtt{Q8}\,\mathtt{R8} := \mathtt{tmp3}^{\mathtt{T}}$

$\mathtt{tmp10} := \mathtt{L1}^{\mathtt{-1}}\,\mathtt{y}$

$\mathtt{tmp11} := \mathtt{Q8}^{\mathtt{T}}\,\mathtt{tmp10}$

$\mathtt{tmp16} := \mathtt{R8}^{\mathtt{-1}}\,\mathtt{tmp11}$

$\mathtt{b} := \mathtt{tmp16}$

1. $LL^T = M$

CL1CK: output algorithms (selected)

$$b = \left(X^T \, L^{-T} \, L^{-1} \, X\right)^{-1} X^T \, L^{-T} \, L^{-1} \, y$$

Algorithm 2

$L1\,L1^T := m$

$tmp3 := x^T\,L1^{-T}$

$Q8\,R8 := tmp3^T$

$tmp10 := L1^{-1}\,y$

$tmp11 := Q8^T\,tmp10$

$tmp16 := R8^{-1}\,tmp11$

$b := tmp16$

1. $LL^T = M$

CL1CK: output algorithms (selected)

$$b = \left(\mathbf{X}^{\mathbf{T}}\,\mathbf{L}^{-\mathbf{T}}\,L^{-1}\,X\right)^{-1}\mathbf{X}^{\mathbf{T}}\,\mathbf{L}^{-\mathbf{T}}\,L^{-1}\,y$$

Algorithm 2

$L1\,L1^{T}$ == m

tmp3 == $x^{T}\,L1^{-T}$

Q8 R8 == $tmp3^{T}$

tmp10 == $L1^{-1}$ y

tmp11 == $Q8^{T}$ tmp10

tmp16 == $R8^{-1}$ tmp11

b == tmp16

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$

CL1CK: output algorithms (selected)

$$b = \left(X^T \, X\right)^{-1} X^T \, L^{-1} \, y$$

Algorithm 2

$L1 \, L1^T := m$

$tmp3 := x^T \, L1^{-T}$

$Q8 \, R8 := tmp3^T$

$tmp10 := L1^{-1} \, y$

$tmp11 := Q8^T \, tmp10$

$tmp16 := R8^{-1} \, tmp11$

$b := tmp16$

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$

CL1CK: output algorithms (selected)

$$b = \left(X^T\ \mathbf{X}\right)^{-1} X^T\ L^{-1}\ y$$

Algorithm 2

```
L1 L1^T := m
tmp3 := x^T L1^-T
Q8 R8 := tmp3^T
tmp10 := L1^-1 y
tmp11 := Q8^T tmp10
tmp16 := R8^-1 tmp11
b := tmp16
```

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$

CL1CK: output algorithms (selected)

$$b = \left(R^T\ Q^T\ Q\ R\right)^{-1} R^T\ Q^T\ L^{-1}\ y$$

Algorithm 2

```
L1 L1^T == m
tmp3 == x^T L1^-T
Q8 R8 == tmp3^T
tmp10 == L1^-1 y
tmp11 == Q8^T tmp10
tmp16 == R8^-1 tmp11
b == tmp16
```

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$

CL1CK: output algorithms (selected)

$$b = \left(R^T\ R\right)^{-1} R^T\ Q^T\ L^{-1}\ y$$

Algorithm 2

$L1\ L1^T := m$

$tmp3 := x^T\ L1^{-T}$

$Q8\ R8 := tmp3^T$

$tmp10 := L1^{-1}\ y$

$tmp11 := Q8^T\ tmp10$

$tmp16 := R8^{-1}\ tmp11$

$b := tmp16$

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$

CL1CK: output algorithms (selected)

$$b = R^{-1}\ R^{-T}\ R^{T}\ Q^{T}\ L^{-1}\ y$$

Algorithm 2

$L1\,L1^{T} == m$

$tmp3 == x^{T}\,L1^{-T}$

$Q8\,R8 == tmp3^{T}$

$tmp10 == L1^{-1}\,y$

$tmp11 == Q8^{T}\,tmp10$

$tmp16 == R8^{-1}\,tmp11$

$b == tmp16$

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$

CL1CK: output algorithms (selected)

$$b = R^{-1}\, Q^T\, L^{-1}\, y$$

Algorithm 2

$L1\, L1^T := m$

$tmp3 := x^T\, L1^{-T}$

$Q8\, R8 := tmp3^T$

$tmp10 := L1^{-1}\, y$

$tmp11 := Q8^T\, tmp10$

$tmp16 := R8^{-1}\, tmp11$

$b := tmp16$

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$

CL1CK: output algorithms (selected)

$$b = R^{-1}\, Q^T\, \mathbf{L}^{-1}\, \mathbf{y}$$

Algorithm 2

```
L1 L1^T == m
tmp3 == x^T L1^-T
Q8 R8 == tmp3^T
tmp10 == L1^-1 y
tmp11 == Q8^T tmp10
tmp16 == R8^-1 tmp11
b == tmp16
```

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$
4. $y \leftarrow L^{-1} y$

CL1CK: output algorithms (selected)

$$b = R^{-1}\, \mathbf{Q^T}\, \mathbf{y}$$

Algorithm 2

```
L1 L1^T == m
tmp3 == x^T L1^-T
Q8 R8 == tmp3^T
tmp10 == L1^-1 y
tmp11 == Q8^T tmp10
tmp16 == R8^-1 tmp11
b == tmp16
```

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$
4. $y \leftarrow L^{-1} y$
5. $b \leftarrow Q^T y$

CL1CK: output algorithms (selected)

$$b = \mathbf{R}^{-1}\,\mathbf{y}$$

Algorithm 2

$L1\,L1^T == m$

$tmp3 == x^T\,L1^{-T}$

$Q8\,R8 == tmp3^T$

$tmp10 == L1^{-1}\,y$

$tmp11 == Q8^T\,tmp10$

$tmp16 == R8^{-1}\,tmp11$

$b == tmp16$

1. $LL^T = M$
2. $X^T \leftarrow X^T L^{-T}$
3. $QR = X$
4. $y \leftarrow L^{-1} y$
5. $b \leftarrow Q^T y$
6. $b \leftarrow R^{-1} b$

CL1CK: output algorithms (selected)

$$b = \left(X^T M^{-1} X\right)^{-1} X^T M^{-1} y$$

Algorithm 8

```
Z1 W1 Z1^T := m
tmp4 := x^T Z1
tmp9 := tmp4 W1^-1
tmp13 := tmp9 tmp4^T
Q20 R20 := tmp13
tmp24 := Z1^T y
tmp30 := tmp9 tmp24
tmp36 := Q20^-1 tmp30
tmp38 := R20^-1 tmp36
b := tmp38
```

1. $ZWZ^T = M$

CL1CK: output algorithms (selected)

$$b = \left(X^T\left(ZWZ^T\right)^{-1}X\right)^{-1}X^T\left(ZWZ^T\right)^{-1}y$$

Algorithm 8

$Z1\,W1\,Z1^T := m$

$tmp4 := x^T\,Z1$

$tmp9 := tmp4\,W1^{-1}$

$tmp13 := tmp9\,tmp4^T$

$Q20\,R20 := tmp13$

$tmp24 := Z1^T\,y$

$tmp30 := tmp9\,tmp24$

$tmp36 := Q20^{-1}\,tmp30$

$tmp38 := R20^{-1}\,tmp36$

$b := tmp38$

1. $ZWZ^T = M$

CL1CK: output algorithms (selected)

$$b = \left(X^T\, Z\, W^{-1}\, Z^T\, X\right)^{-1} X^T\, Z\, W^{-1}\, Z^T\, y$$

Algorithm 8

```
Z1 W1 Z1^T == m
tmp4 == x^T Z1
tmp9 == tmp4 W1^-1
tmp13 == tmp9 tmp4^T
Q20 R20 == tmp13
tmp24 == Z1^T y
tmp30 == tmp9 tmp24
tmp36 == Q20^-1 tmp30
tmp38 == R20^-1 tmp36
b == tmp38
```

1. $ZWZ^T = M$
2. ...
3. ...
4. ...

1) Knowledge encoded

1) Knowledge encoded

- Operand types: scalars, vectors, matrices

  Size: $(1|m) \times (1|n)$

  Properties: full rank, orthogonal, symmetric, triangular, SPD, diagonal, identity, ...

1) Knowledge encoded

- Operand types: scalars, vectors, matrices

  Size: $(1|m) \times (1|n)$

  Properties: full rank, orthogonal, symmetric, triangular, SPD, diagonal, identity, ...

- Linear Algebra operators: +, -, $\star$, ^{T}, $^{-1}$

  Linear Algebra properties: precedence, associativity, commutativity, distributivity of transposition and inversion...

1) Knowledge encoded

- Operand types: scalars, vectors, matrices

  Size: $(1|m) \times (1|n)$

  Properties: full rank, orthogonal, symmetric, triangular, SPD, diagonal, identity, ...

- Linear Algebra operators: +, -, $\star$, ^{T}, $^{-1}$

  Linear Algebra properties: precedence, associativity, commutativity, distributivity of transposition and inversion...

- BLAS' and LAPACK's interface

1) Knowledge encoded

- Operand types: scalars, vectors, matrices

  Size: $(1|m) \times (1|n)$

  Properties: full rank, orthogonal, symmetric, triangular, SPD, diagonal, identity, ...

- Linear Algebra operators: +, -, $\star$, ^{T}, $^{-1}$

  Linear Algebra properties: precedence, associativity, commutativity, distributivity of transposition and inversion...

- BLAS' and LAPACK's interface

- Guidelines:
  - inversion $\rightarrow$ factorization to triangular/diagonal matrices
  - avoid redundant computations
    common expression substitution
  - expression decomposed into basic tasks (BLAS/LAPACK)

2) Knowledge derived

2) Knowledge derived

- Inputs/outputs
  X, M are inputs, b is output
  $\Rightarrow \quad (X^T M^{-1} X)^{-1}$ is known, Xb is unknown

2) Knowledge derived

- Inputs/outputs
  X, M are inputs, b is output
  $\Rightarrow \quad (X^T M^{-1} X)^{-1}$ is known, Xb is unknown

- Operand type and size
  $X \in \mathcal{R}^{n \times p}, \; M \in \mathcal{R}^{n \times n} \quad \Rightarrow \quad X^T M^{-1} X \in \mathcal{R}^{p \times p}$

2) Knowledge derived

- Inputs/outputs
  X, M are inputs, b is output
  $\Rightarrow \quad (X^T M^{-1} X)^{-1}$ is known, Xb is unknown

- Operand type and size
  $X \in \mathcal{R}^{n \times p},\ M \in \mathcal{R}^{n \times n} \quad \Rightarrow \quad X^T M^{-1} X \in \mathcal{R}^{p \times p}$

- Properties
  - L is lower tri., D is diagonal, U is upper tri.
    $\Rightarrow \quad LD^{-1}U^{-T}$ is lower triangular
  - X is full rank, $SPD(M)$
    $\Rightarrow \quad SPD(X^T M^{-1} X)$

2) Knowledge derived

- Inputs/outputs
  X, M are inputs, b is output
  $\Rightarrow \quad (X^T M^{-1} X)^{-1}$ is known, Xb is unknown

- Operand type and size
  $X \in \mathcal{R}^{n \times p}, \ M \in \mathcal{R}^{n \times n} \quad \Rightarrow \quad X^T M^{-1} X \in \mathcal{R}^{p \times p}$

- Properties
  - L is lower tri., D is diagonal, U is upper tri.
    $\Rightarrow \quad LD^{-1}U^{-T}$ is lower triangular
  - X is full rank, $SPD(M)$
    $\Rightarrow \quad SPD(X^T M^{-1} X)$

- Arithmetic
  $(R^T Q^T Q R)^{-1} R^T Q^T \quad \Rightarrow \quad R^{-1} Q^T$

Experiments

The problem: the real setup

Compute $B \in \mathcal{R}^{p \times m \times t}$

$$b_{ij} = \left(X_i^T M_j^{-1} X_i\right)^{-1} X_i^T M_j^{-1} y_j$$

$M \in \mathcal{R}^{n \times n}$,	with $n \in [10^3, \ldots, 10^4]$
$X \in \mathcal{R}^{n \times p}$,	with $p \in [1, \ldots, 20]$

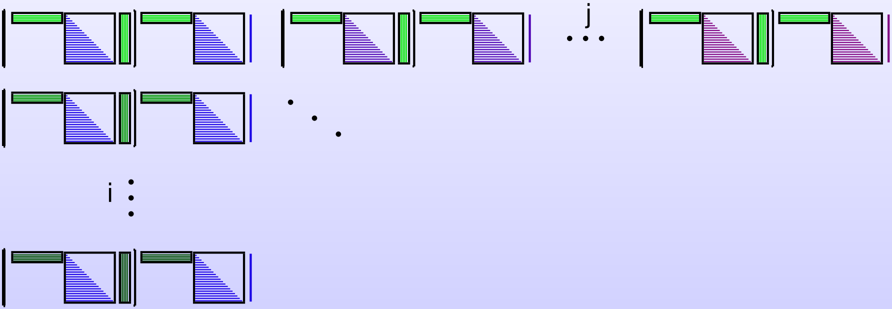
and

$i = 1 \ldots m$,	with $m \approx 10^6 - 10^7$
$j = 1 \ldots t$,	where t is either 1 or $\approx 10^5$

$\Rightarrow$ Terabytes of data!

Also, $M_j = h_j \phi + (1 - h_j) I$

2D sequence of problems



Naive approach: `for i, for j, ...`

$$b_{ij} = \left(X_i^T M_j^{-1} X_i\right)^{-1} X_i^T M_j^{-1} y_j$$

$$
\begin{array}{l}
\texttt{for}\ \ i = 1 : m \\
\quad \texttt{for}\ \ j = 1 : t \\
\qquad LL^T = M_j \\
\qquad X^T \leftarrow X_i^T L^{-T} \\
\qquad QR = X \\
\qquad y \leftarrow L^{-1} y_j \\
\qquad b \leftarrow Q^T y \\
\qquad b_{ij} \leftarrow R^{-1} b
\end{array}
$$

Tracking the dependencies

$$LL^T = M_j \quad \Rightarrow \quad L_j L_j^T = M_j$$

$$X^T \leftarrow X_i^T L_j^{-T} \quad \Rightarrow \quad X_{ij}^T \leftarrow X_i^T L_j^{-T}$$

```
for i = 1 : m
    for j = 1 : t
```

$$L_j L_j^T = M_j$$

$$X_{ij}^T \leftarrow X_i^T L_j^{-T}$$

$$Q_{ij} R_{ij} = X_{ij}$$

$$y_j \leftarrow L_j^{-1} y_j$$

$$b_{ij} \leftarrow Q_{ij}^T y_j$$

$$b_{ij} \leftarrow R_{ij}^{-1} b_{ij}$$

Loop Transposition

$$LL^T = M_j \quad \Rightarrow \quad L_j L_j^T = M_j$$

$$X^T \leftarrow X_i^T L_j^{-T} \quad \Rightarrow \quad X_{ij}^T \leftarrow X_i^T L_j^{-T}$$

```
for i = 1 : m
    for j = 1 : t
        L_j L_j^T = M_j
        X_ij^T ← X_i^T L_j^{-T}
        Q_ij R_ij = X_ij
        y_j ← L_j^{-1} y_j
        b_ij ← Q_ij^T y_j
        b_ij ← R_ij^{-1} b_ij
```

```
for j = 1 : t
    for i = 1 : m
        L_j L_j^T = M_j
        X_ij^T ← X_i^T L_j^{-T}
        Q_ij R_ij = X_ij
        y_j ← L_j^{-1} y_j
        b_ij ← Q_ij^T y_j
        b_ij ← R_ij^{-1} b_ij
```

Reordering

$$LL^T = M_j \quad \Rightarrow \quad L_j L_j^T = M_j$$

$$X^T \leftarrow X_i^T L_j^{-T} \quad \Rightarrow \quad X_{ij}^T \leftarrow X_i^T L_j^{-T}$$

for $i = 1 : m$
  for $j = 1 : t$
    $L_j L_j^T = M_j$
    $X_{ij}^T \leftarrow X_i^T L_j^{-T}$
    $Q_{ij} R_{ij} = X_{ij}$
    $y_j \leftarrow L_j^{-1} y_j$
    $b_{ij} \leftarrow Q_{ij}^T y_j$
    $b_{ij} \leftarrow R_{ij}^{-1} b_{ij}$

for $j = 1 : t$
  for $i = 1 : m$
    $L_j L_j^T = M_j$
    $X_{ij}^T \leftarrow X_i^T L_j^{-T}$
    $Q_{ij} R_{ij} = X_{ij}$
    $y_j \leftarrow L_j^{-1} y_j$
    $b_{ij} \leftarrow Q_{ij}^T y_j$
    $b_{ij} \leftarrow R_{ij}^{-1} b_{ij}$

Reordering

$$LL^T = M_j \quad \Rightarrow \quad L_j L_j^T = M_j$$

$$X^T \leftarrow X_i^T L_j^{-T} \quad \Rightarrow \quad X_{ij}^T \leftarrow X_i^T L_j^{-T}$$

$\texttt{for } i = 1:m$
- $\texttt{for } j = 1:t$
  - $L_j L_j^T = M_j$
  - $X_{ij}^T \leftarrow X_i^T L_j^{-T}$
  - $Q_{ij} R_{ij} = X_{ij}$
  - $y_j \leftarrow L_j^{-1} y_j$
  - $b_{ij} \leftarrow Q_{ij}^T y_j$
  - $b_{ij} \leftarrow R_{ij}^{-1} b_{ij}$

$\texttt{for } j = 1:t$
- $L_j L_j^T = M_j$
- $y_j \leftarrow L_j^{-1} y_j$
- $\texttt{for } i = 1:m$
  - $X_{ij}^T \leftarrow X_i^T L_j^{-T}$
  - $Q_{ij} R_{ij} = X_{ij}$
  - $b_{ij} \leftarrow Q_{ij}^T y_j$
  - $b_{ij} \leftarrow R_{ij}^{-1} b_{ij}$

Cl1ck ver. 2.0

Input

```
precond = {
  {X, {"Input", "Matrix", "FullRank"}, { i } }
  {y, {"Input", "Vector" }, { j } }
  {Phi, {"Input", "Matrix", "Symmetric", "Lower"}, { } }
  {h, {"Input", "Scalar" }, { i } }
  {b, {"Output", "Vector" }, { i, j } }
};
addProperties = {
  plus[times[h,Phi], times[plus[1,minus[h]], Id],
  { "SPD" } };
```

- Dependencies tracking
- Loop transposition
- Reordering

Algorithms and code

Algorithm 2

$L1\,L1^T == m$

$tmp3 == x^T\,L1^{-T}$

$Q8\,R8 == tmp3^T$

$tmp10 == L1^{-1}\,y$

$tmp11 == Q8^T\,tmp10$

$tmp16 == R8^{-1}\,tmp11$

$b == tmp16$

```
function [b] = FGLS_2_1(x, y, m, sn, sp, nms, nxs)
   b = zeros(sp, nxs * nms);
   for i = 1:nxs
      x_i = x(:, sp*(i-1)+1 : sp*i);
      for j = 1:nms
         y_j = y(:, j);
         m_j = m(:, sn*(j-1)+1 : sn*j);
         L1 = chol( m_j, 'lower' );
         T_1 = x_i' / L1';
         T_2 = L1 \ y_j;
         [Q16, R16] = qr( T_1' );
         T_3 = Q16' * T_2;
         T_4 = R16 \ T_3;
         b(:, i + (j-1)*nxs) = T_4;
      end
   end
end
```

Algorithms and code

Algorithm 2

$L1\,L1^T == m$

$tmp3 == x^T\,L1^{-T}$

$Q8\,R8 == tmp3^T$

$tmp10 == L1^{-1}\,y$

$tmp11 == Q8^T\,tmp10$

$tmp16 == R8^{-1}\,tmp11$

$b == tmp16$

```
function [b] = FGLS_2_2(x, y, m, sn, sp, nms, nxs)
   b = zeros(sp, nxs * nms);
   for j = 1:nms
      y_j = y(:, j);
      m_j = m(:, sn*(j-1)+1 : sn*j);
      L1 = chol( m_j, 'lower' );
      T_2 = L1 \ y_j;
      for i = 1:nxs
         x_i = x(:, sp*(i-1)+1 : sp*i);
         T_1 = x_i' / L1';
         [Q16, R16] = qr( T_1' );
         T_3 = Q16' * T_2;
         T_4 = R16 \ T_3;
         b(:, i + (j-1)*nxs) = T_4;
      end
   end
end
```

Algorithms and code

Algorithm 18

$$Z1\ W1\ Z1^T == phi$$
$$Z1\ Z1^T == id$$
$$tmp2 == -(h\ id) + 1\ id + h\ W1$$
$$tmp7 == x^T\ Z1$$
$$D3\ D3^T == tmp2$$
$$tmp22 == tmp7\ D3^{-T}$$
$$tmp31 == tmp22\ tmp22^T$$
$$L6\ L6^T == tmp31$$
$$tmp70 == Z1^T\ y$$
$$tmp98 == D3^{-1}\ tmp70$$
$$tmp122 == tmp22\ tmp98$$
$$tmp148 == L6^{-1}\ tmp122$$
$$tmp165 == L6^{-T}\ tmp148$$
$$b == tmp165$$

Algorithms and code

Algorithm 18

```
Z1 W1 Z1^T == phi
Z1 Z1^T == id
tmp2 == - (h id) + 1 id + h W1
tmp7 == x^T Z1
D3 D3^T == tmp2
tmp22 == tmp7 D3^-T
tmp31 == tmp22 tmp22^T
L6 L6^T == tmp31
tmp70 == Z1^T y
tmp98 == D3^-1 tmp70
tmp122 == tmp22 tmp98
tmp148 == L6^-1 tmp122
tmp165 == L6^-T tmp148
b == tmp165
```

$$M = h_j \Phi + (1 - h_j) I$$

$$ZWZ^T = \Phi, \quad ZZ^T = I$$

$$\Rightarrow$$

$$M^{-1} = (h_j ZWZ^T + (1 - h_j) ZZ^T)^{-1} =$$

$$Z \left(h_j W + (1 - h_j I)\right)^{-1} Z^T$$

Algorithms and code

Algorithm 18

$Z1\, W1\, Z1^T == phi$

$Z1\, Z1^T == id$

$tmp2 == -(h\, id) + 1\, id + h\, W1$

$tmp7 == x^T\, Z1$

$D3\, D3^T == tmp2$

$tmp22 == tmp7\, D3^{-T}$

$tmp31 == tmp22\, tmp22^T$

$L6\, L6^T == tmp31$

$tmp70 == Z1^T\, y$

$tmp98 == D3^{-1}\, tmp70$

$tmp122 == tmp22\, tmp98$

$tmp148 == L6^{-1}\, tmp122$

$tmp165 == L6^{-T}\, tmp148$

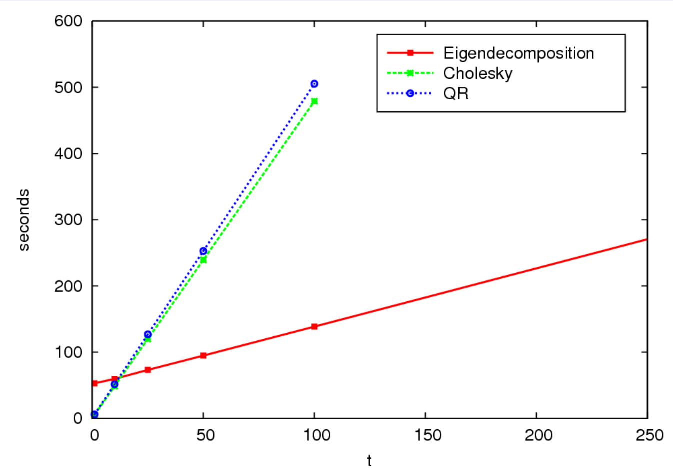
$b == tmp165$

```
function [b] = FGLS_18_2(x, y, phi, h, sn, sp, nxs, nys)
   b = zeros(sp, nxs * nys);
   [Z1, W1] = eig( phi );
   for j = 1:nys
      y_j = y(:, j);
      h_j = h(j);
      T_1 = - (h_j * eye(sn)) + 1 * eye(sn) + h_j * W1;
      D3 = diag( sqrt( diag( T_1 ) ) );
      T_5 = Z1' * y_j;
      T_6 = D3 \ T_5;
      for i = 1:nxs
         x_i = x(:, sp*(i-1)+1 : sp*i);
         T_2 = x_i' * Z1;
         T_3 = T_2 / D3';
         T_4 = T_3 * T_3';
         L6 = chol( T_4, 'lower' );
         T_7 = T_3 * T_6;
         T_8 = L6 \ T_7;
         T_9 = L6' \ T_8;
         b(:, i + (j-1)*nxs) = T_9;
      end
   end
end
```

Performance

C + LAPACK/BLAS

$$n = 5000 \quad p = 4 \quad m = 3000$$



Conclusions & future work

CL1CK

- Automated system: algorithm & code generator
- Knowledge: encoded and inferred
- n-fold speedups

Conclusions & future work

CL1CK

- Automated system: algorithm & code generator
- Knowledge: encoded and inferred
- n-fold speedups

TODO list

- **Cost Analysis**   stability analysis
- Mixed models: out-of-core; sparsity
- CL1CK: extension to tensors, derivatives

Deutsche Forschungsgemeinschaft
DFG

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.