Das Verfahren von Cholesky

—

ein geodätischer Beitrag zur numerischen Mathematik

Wolf-Dieter Schuh

Institut für Geodäsie und Geoinformation
Universität Bonn

Geodätische Woche 2014

Session 6: Theoretische Geodäsie

Berlin, 7.-9. Oktober 2014

Contents

André-Louis Cholesky

(1875 - 1918)

Commandant Benoit (1924):

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'EQUATIONS LINÉAIRES.

Bulletin géodésique, 1924, 2, 67-77

I. — NOTICES SCIENTIFIQUES

Commandant BENOIT[1].

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'ÉQUATIONS LINÉAIRES.

(Procédé du Commandant CHOLESKY[2].)

Le Commandant d'Artillerie Cholesky, du Service géographique de l'Armée, tué pendant la grande guerre, a imaginé, au cours de recherches sur la compensation des réseaux géodésiques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des moindres carrés à des équations linéaires en nombre inférieur à celui des inconnues. Il en a conclu une méthode générale de résolution des équations linéaires.

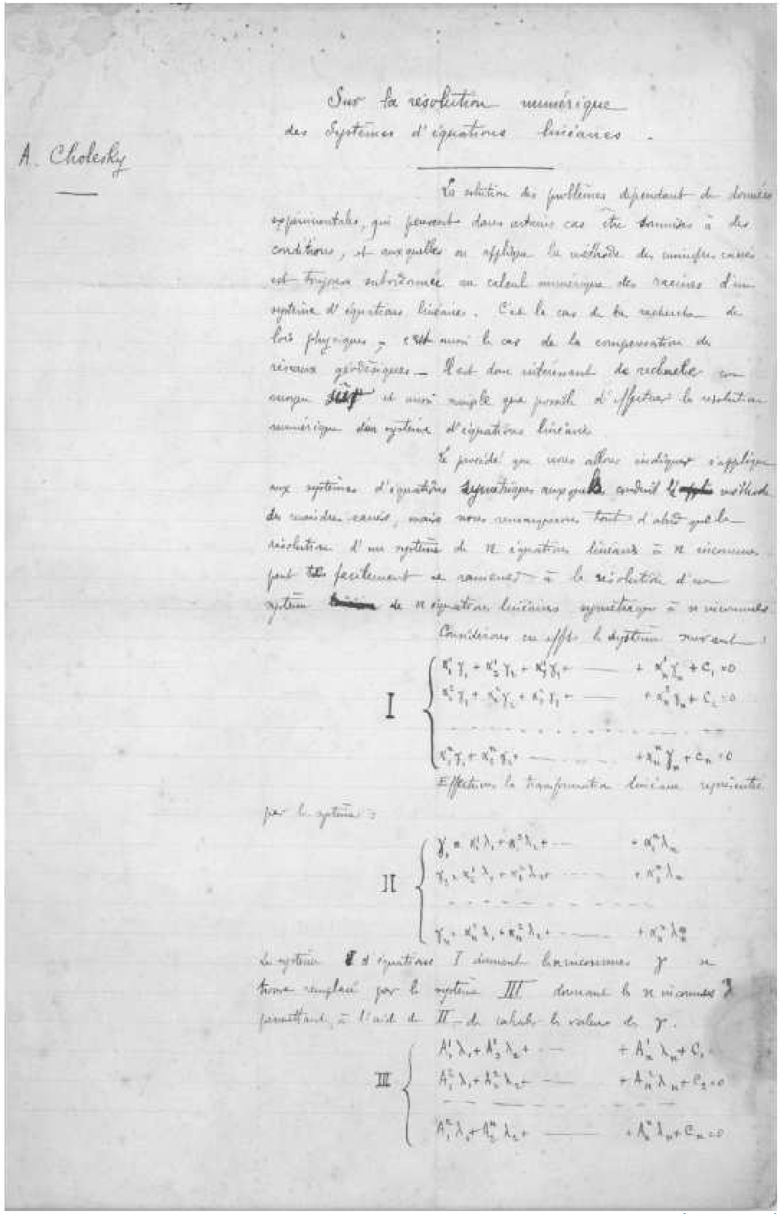
Nous suivrons, pour la démonstration de cette méthode, la progression même qui a servi au Commandant Cholesky pour l'imaginer.

1. De l'Artillerie coloniale, ancien officier géodésien au Service géographique de l'Armée et au Service géographique de l'Indo-Chine, Membre du Comité national français de Géodésie et Géophysique.

2. Sur le Commandant Cholesky, tué à l'ennemi le 31 août 1918, voir la notice biographique insérée dans le volume du *Bulletin géodésique* de 1922 intitulé : *Union géodésique et géophysique internationale, Première Assemblée générale, Rome, mai 1922, Section de Géodésie*, Toulouse, Privat, 1922, in-8°, 241 p., pp. 159 à 161.

5

André-Louis Cholesky



Sur la résolution numérique des Systèmes d'équations linéaires

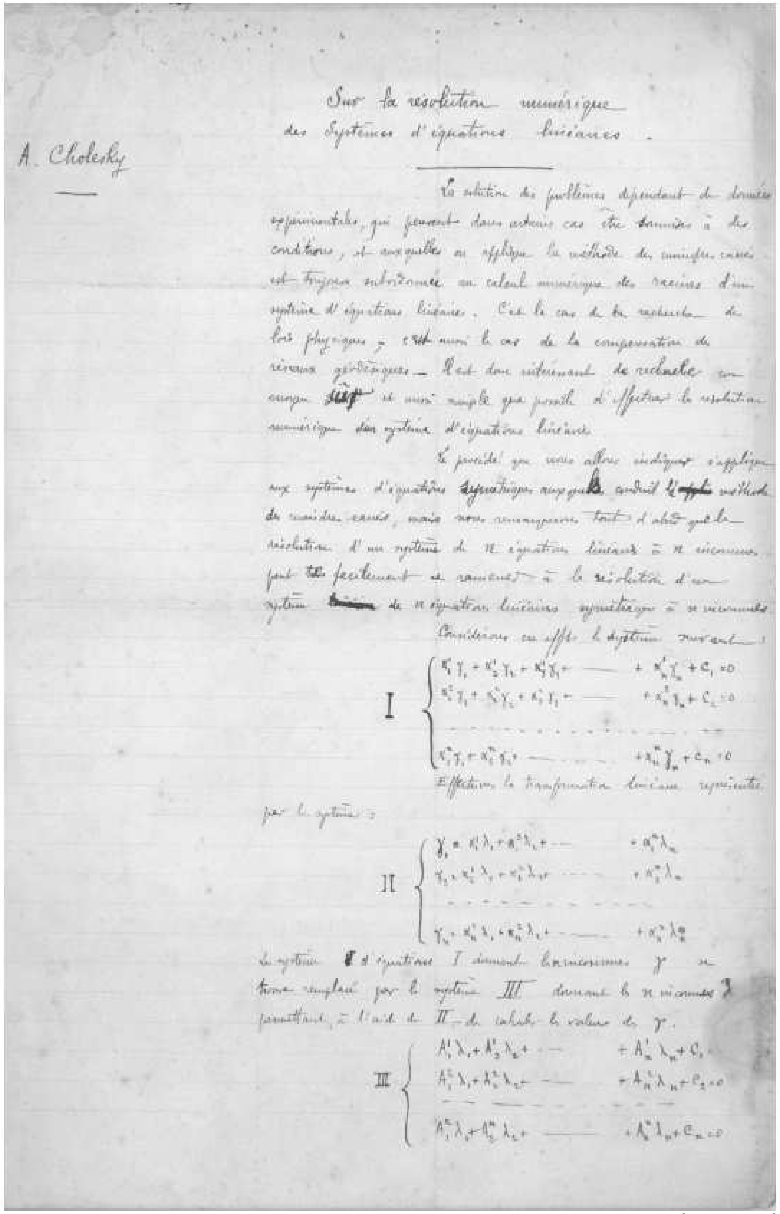
A. Cholesky

Brezinski (2006?)

unpublished manuscript:

Cholesky, A. (2. Dec. 1910): *Sur la résolution numérique des systèmes d' équations linéaires* (On the numerical solution of systems of linear equations)

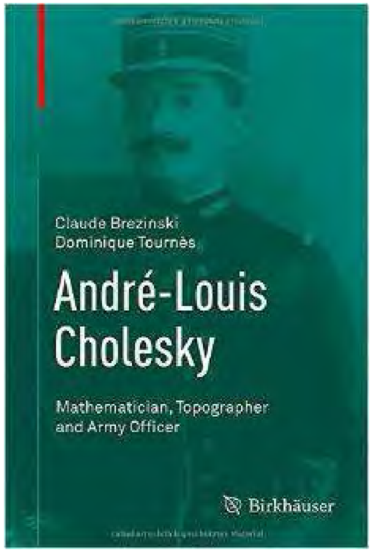
André-Louis Cholesky



Brezinski (2006?)

unpublished manuscript:

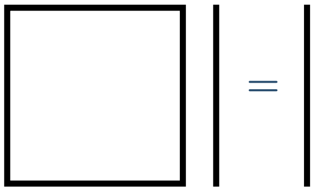
Cholesky, A. (2. Dec. 1910):
Sur la résolution numérique des systèmes d' équations linéaires
(On the numerical solution of systems of linear equations)



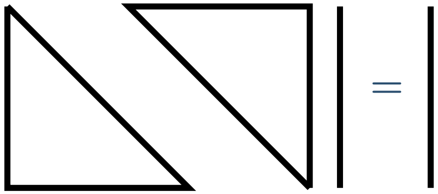
(Sept. 2014)

Cholesky approach

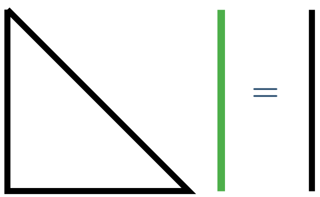
$$\boldsymbol{N}\boldsymbol{x} = \boldsymbol{n}$$

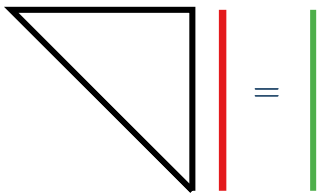


$$\boldsymbol{R}^T \underbrace{\boldsymbol{R}\,\boldsymbol{x}}_{\boldsymbol{z}} = \boldsymbol{n}$$



$$\boldsymbol{R}^T \boldsymbol{z} = \boldsymbol{n} \qquad \boldsymbol{R}\,\boldsymbol{x} = \boldsymbol{z}$$





forward substitution backward substitution

$\Rightarrow \boldsymbol{z}$ $\Rightarrow \boldsymbol{x}$

Cholesky factorization

$$\boldsymbol{N} = \boldsymbol{R}^T \boldsymbol{R}$$

$$\begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{12} & n_{22} & n_{23} \\ n_{13} & n_{23} & n_{33} \end{bmatrix} = \begin{bmatrix} r_{11} & & \\ r_{12} & r_{22} & \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ & r_{22} & r_{23} \\ & & r_{33} \end{bmatrix}$$

$$r_{ii} = \sqrt{n_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}\,, \qquad i=1,\ldots,m$$

$$r_{ij} = \left(n_{ij} - \sum_{k=1}^{i-1} r_{ki} r_{kj} \right) / r_{ii}\,, \qquad i=1,\ldots,m \quad j=i+1,\ldots,m$$

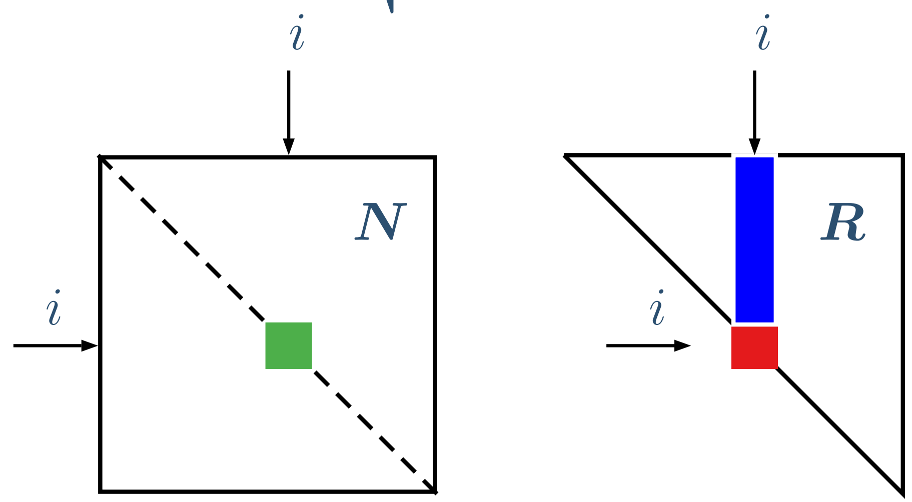
Cholesky (forward) reduction

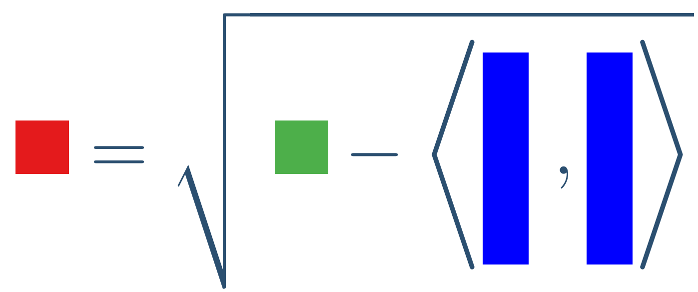
Equation interpretation/reading:

$$r_{ii} = \sqrt{n_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}$$

Equation interpretation/reading:

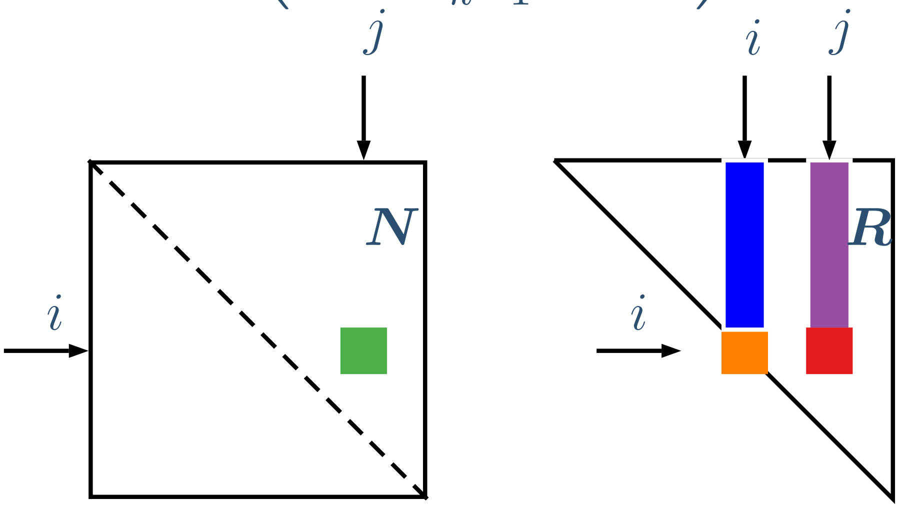
$$r_{ii} = \sqrt{n_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}$$

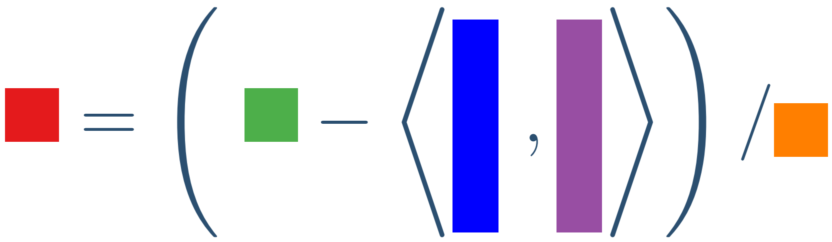




Equation interpretation/reading:

$$r_{ij} = \left(n_{ij} - \sum_{k=1}^{i-1} r_{ki} r_{kj} \right) / r_{ii}$$





Cholesky inversion

$$\boldsymbol{N} \Rightarrow \boldsymbol{R} \Rightarrow \boldsymbol{N}^{-1}$$

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ & r_{22} & r_{23} \\ & & r_{33} \end{bmatrix} \Rightarrow \begin{bmatrix} n_{11}^{(-1)} & n_{12}^{(-1)} & n_{13}^{(-1)} \\ n_{12}^{(-1)} & n_{22}^{(-1)} & n_{23}^{(-1)} \\ n_{13}^{(-1)} & n_{23}^{(-1)} & n_{33}^{(-1)} \end{bmatrix}$$

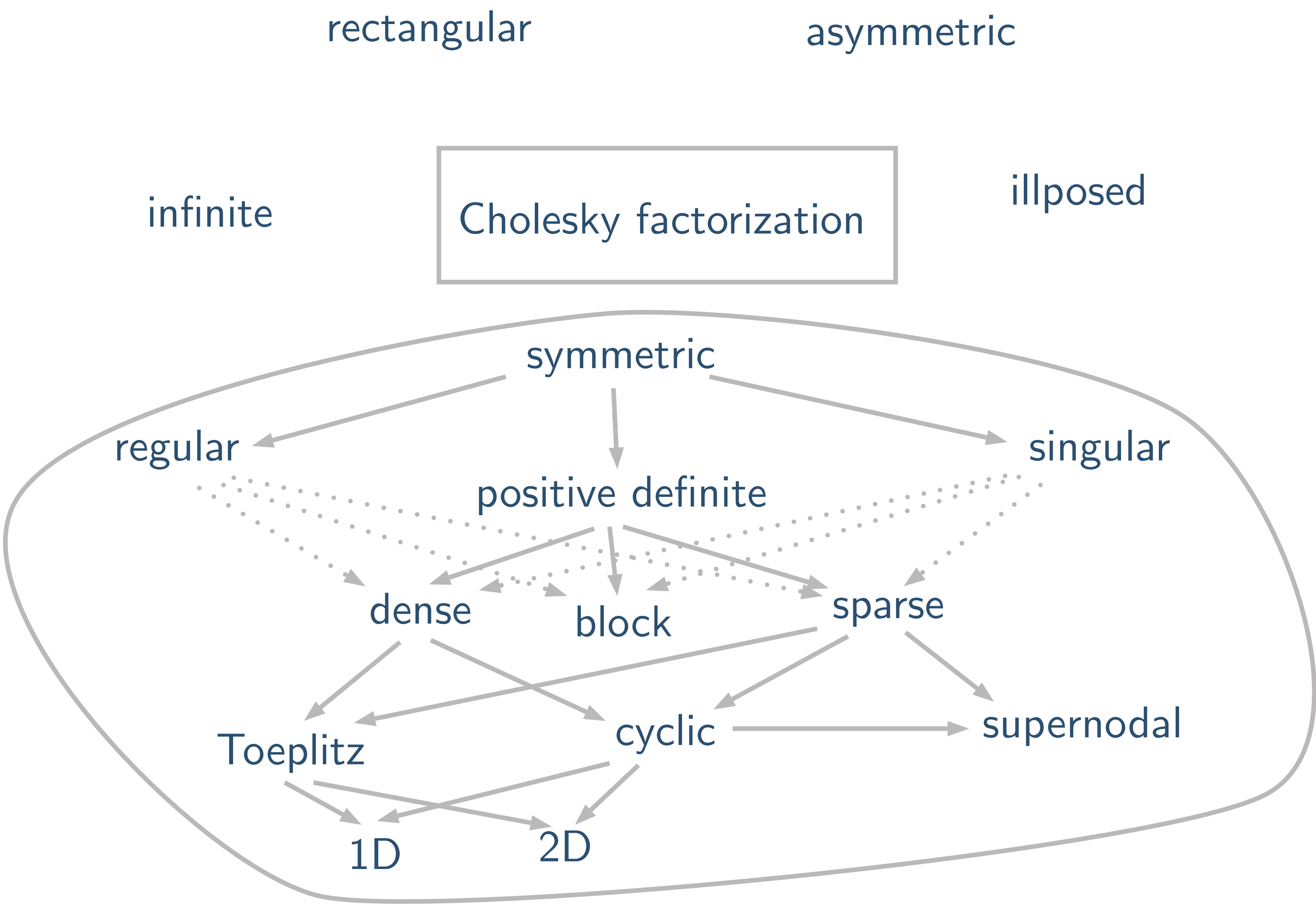
$$n_{ii}^{(-1)} = \frac{1}{r_{ii}^2} - \frac{1}{r_{ii}} \sum_{k=i+1}^{m} r_{ik} n_{ik}^{(-1)}\,, \qquad i=1,\ldots,m$$

$$n_{ij}^{(-1)} = - \sum_{k=i+1}^{m} r_{ik} n_{kj}^{(-1)}\,, \qquad i=1,\ldots,m \quad j=i+1,\ldots,m$$

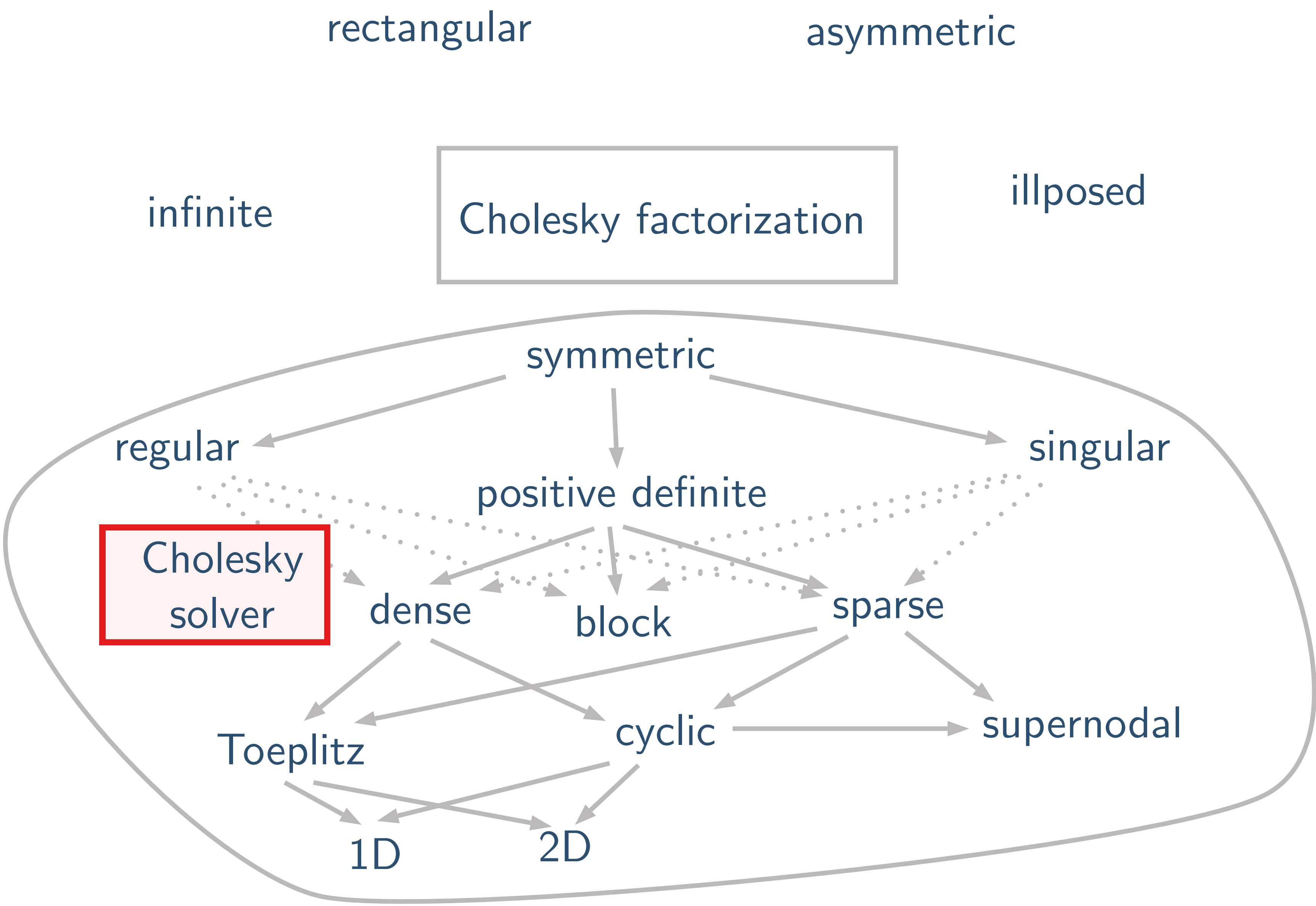
Cholesky inversion by backward substitution

Cholesky applications

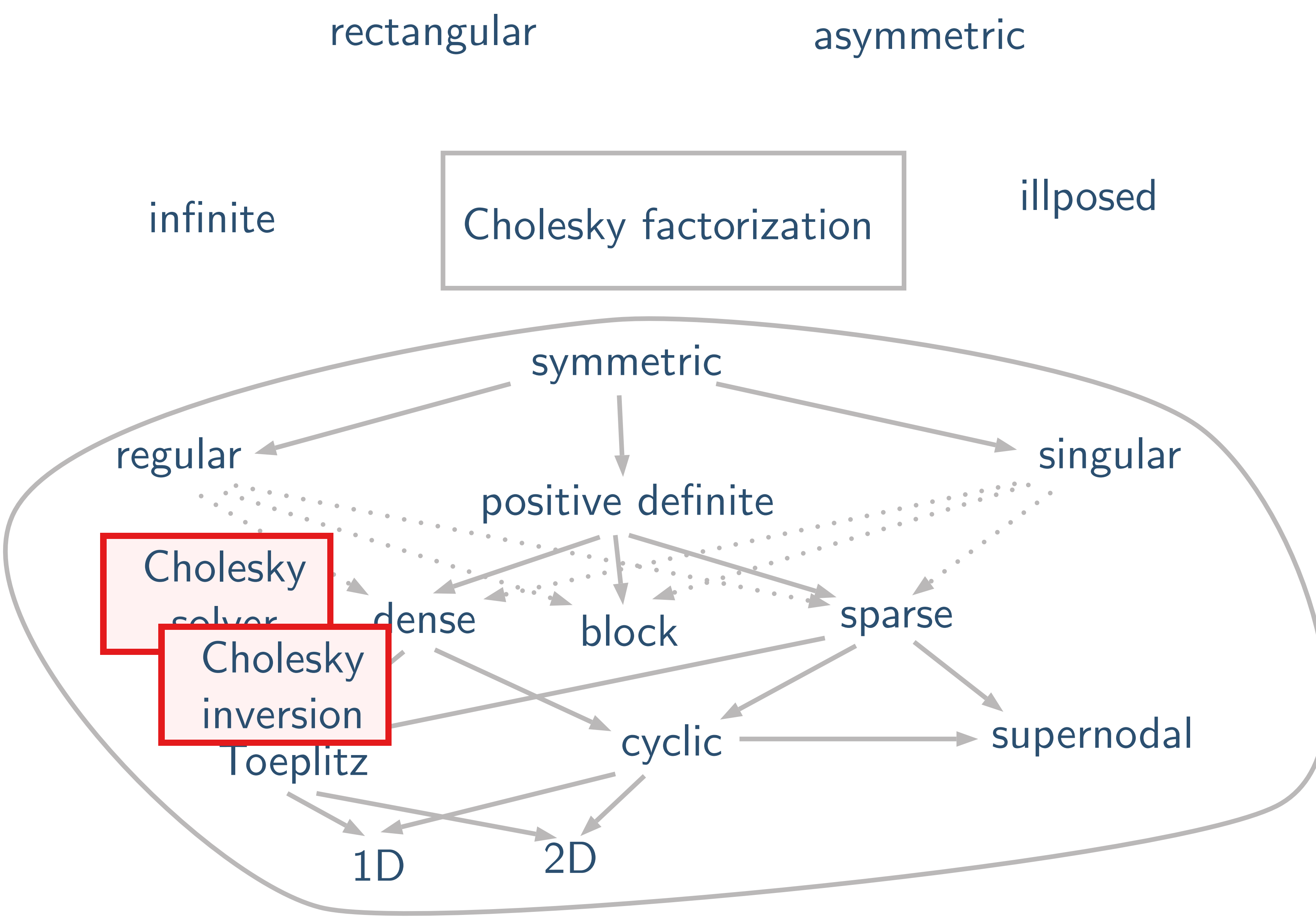
Cholesky application



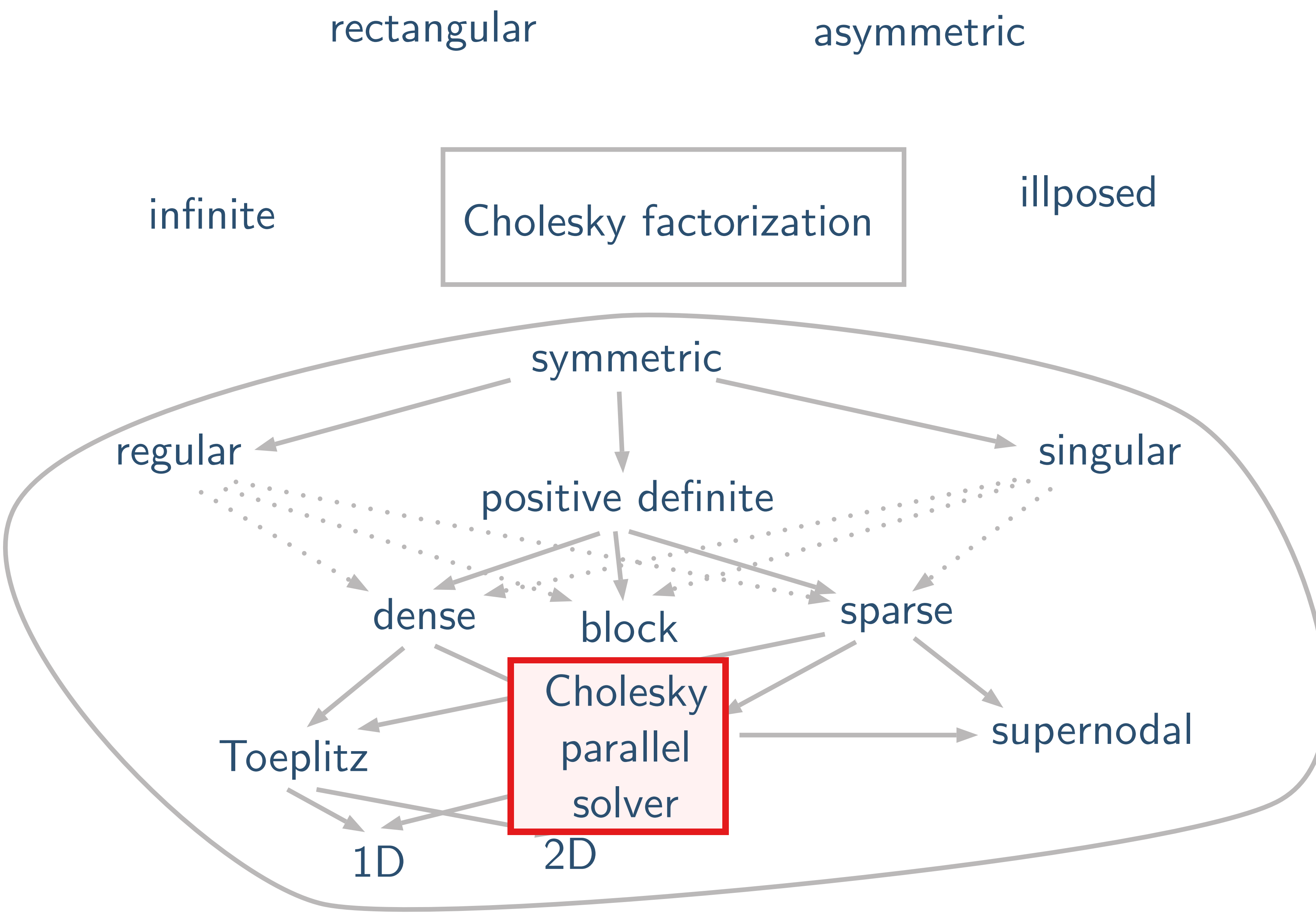
Cholesky application



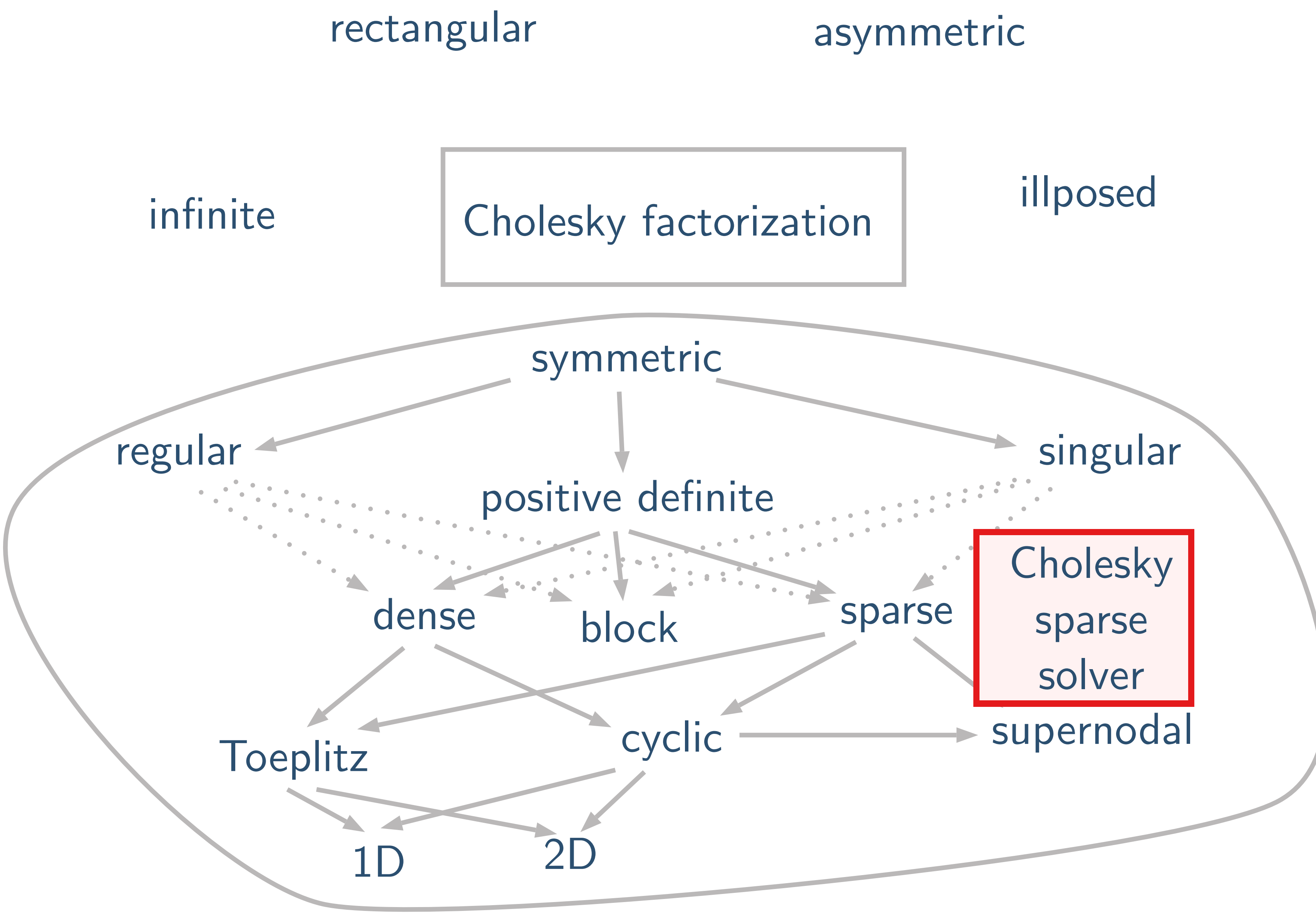
Cholesky application



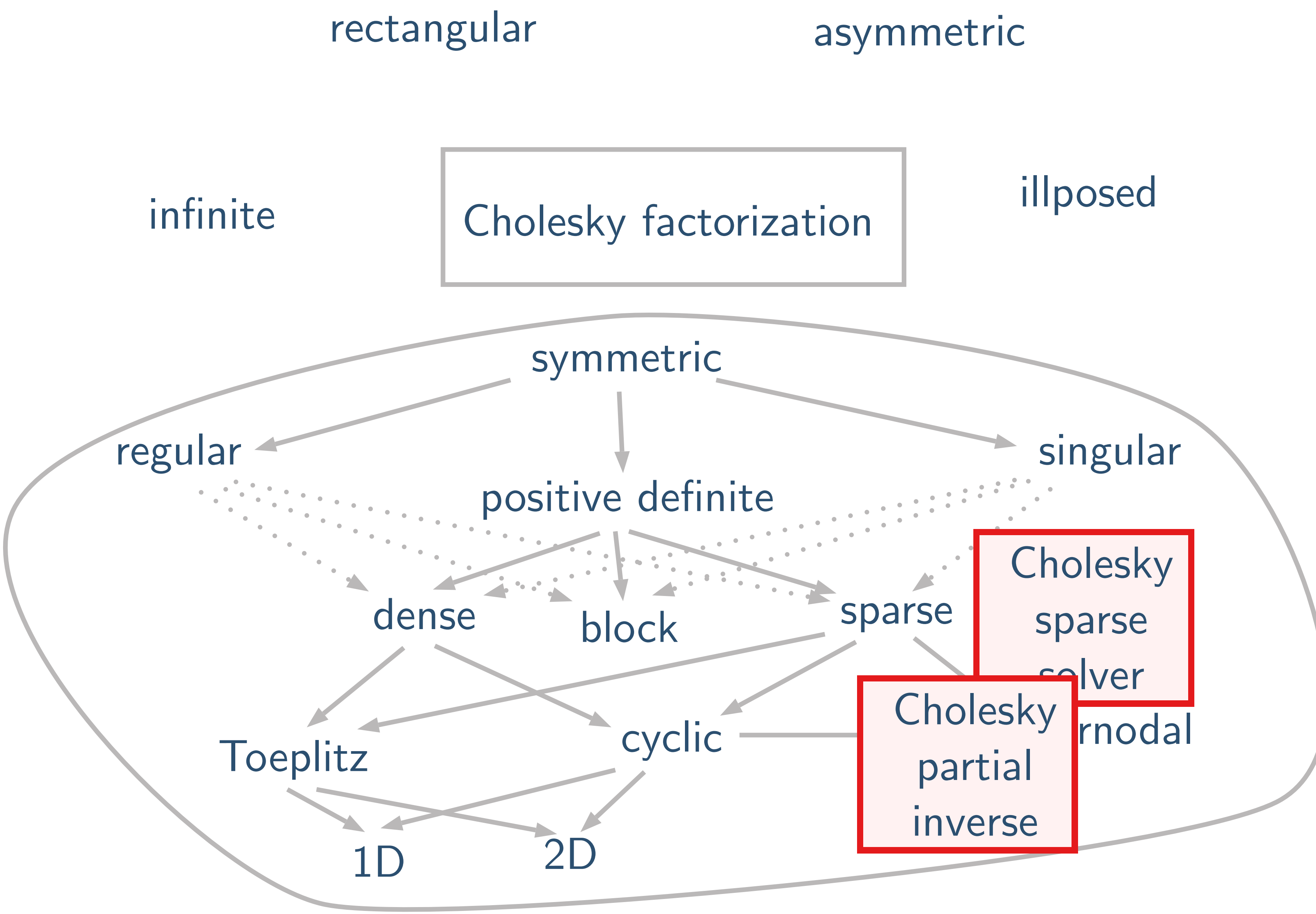
Cholesky application



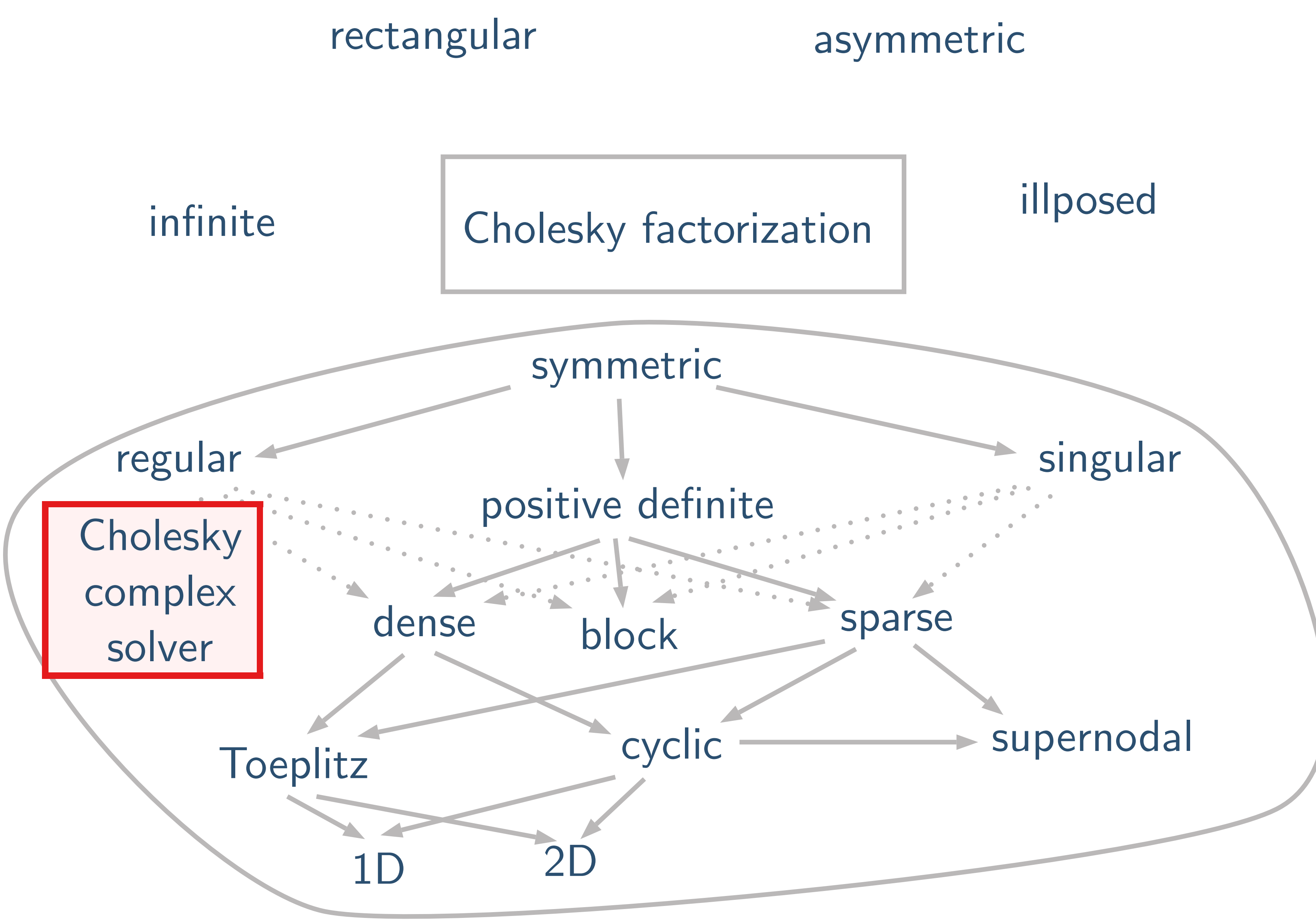
Cholesky application



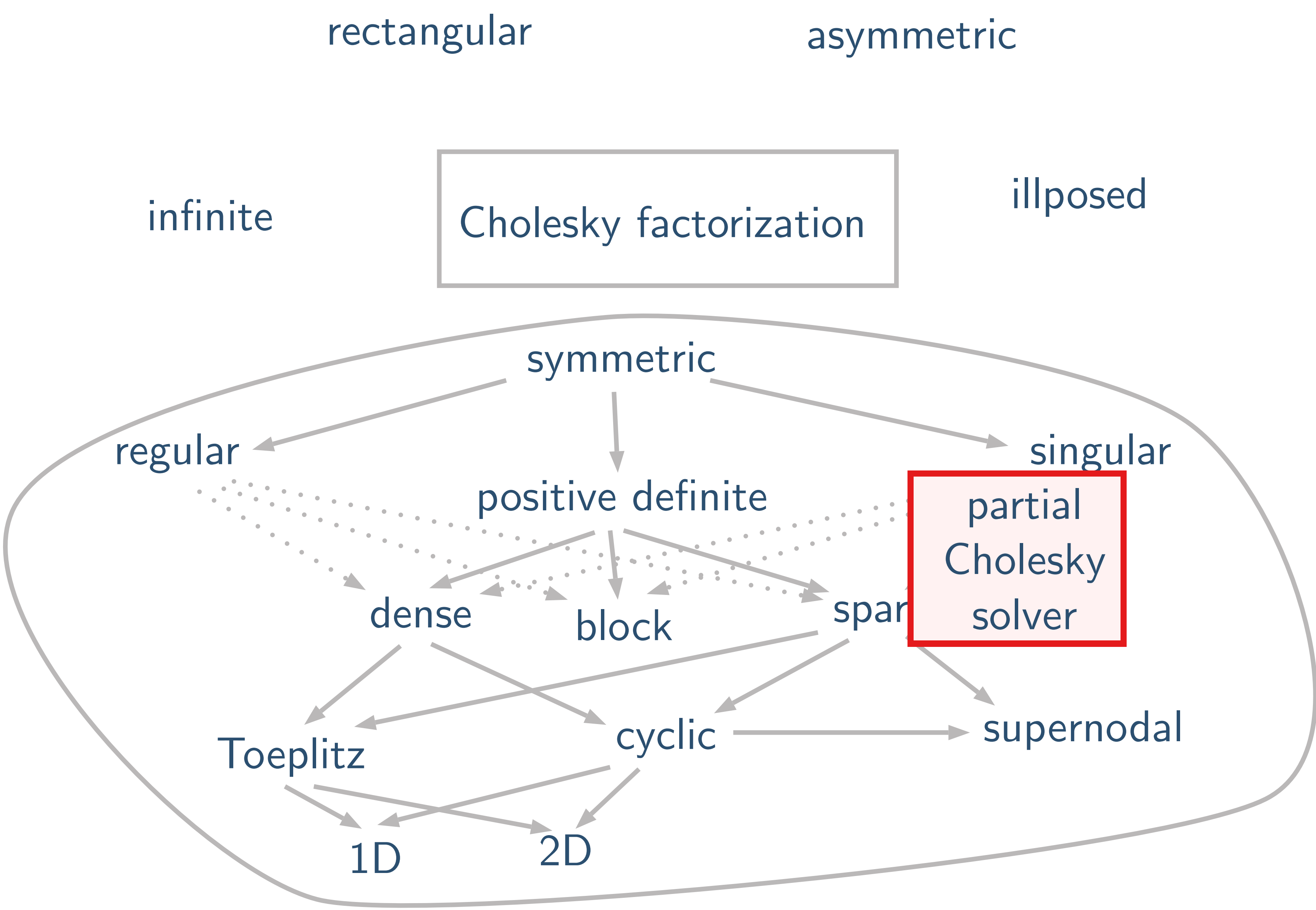
Cholesky application



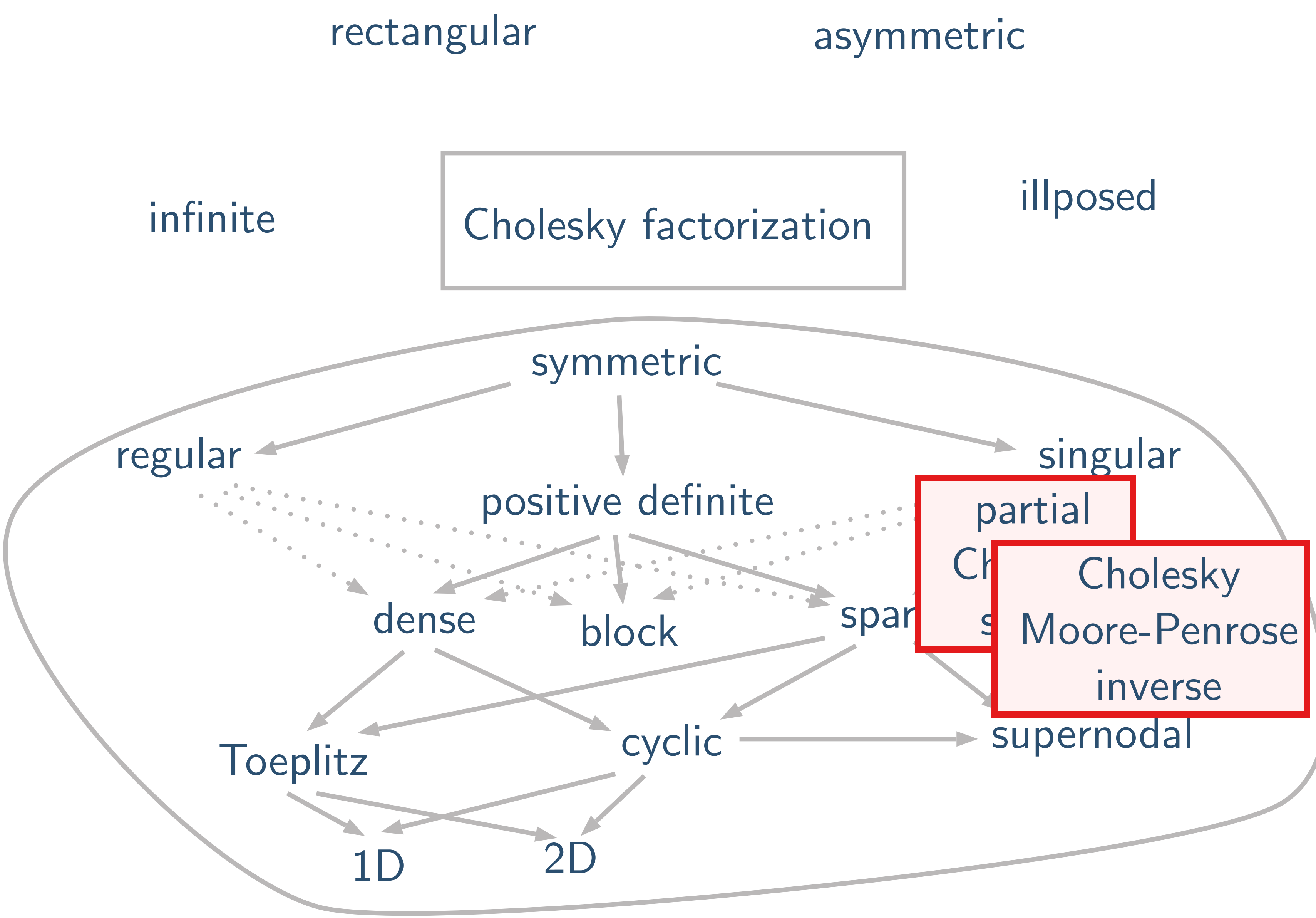
Cholesky application



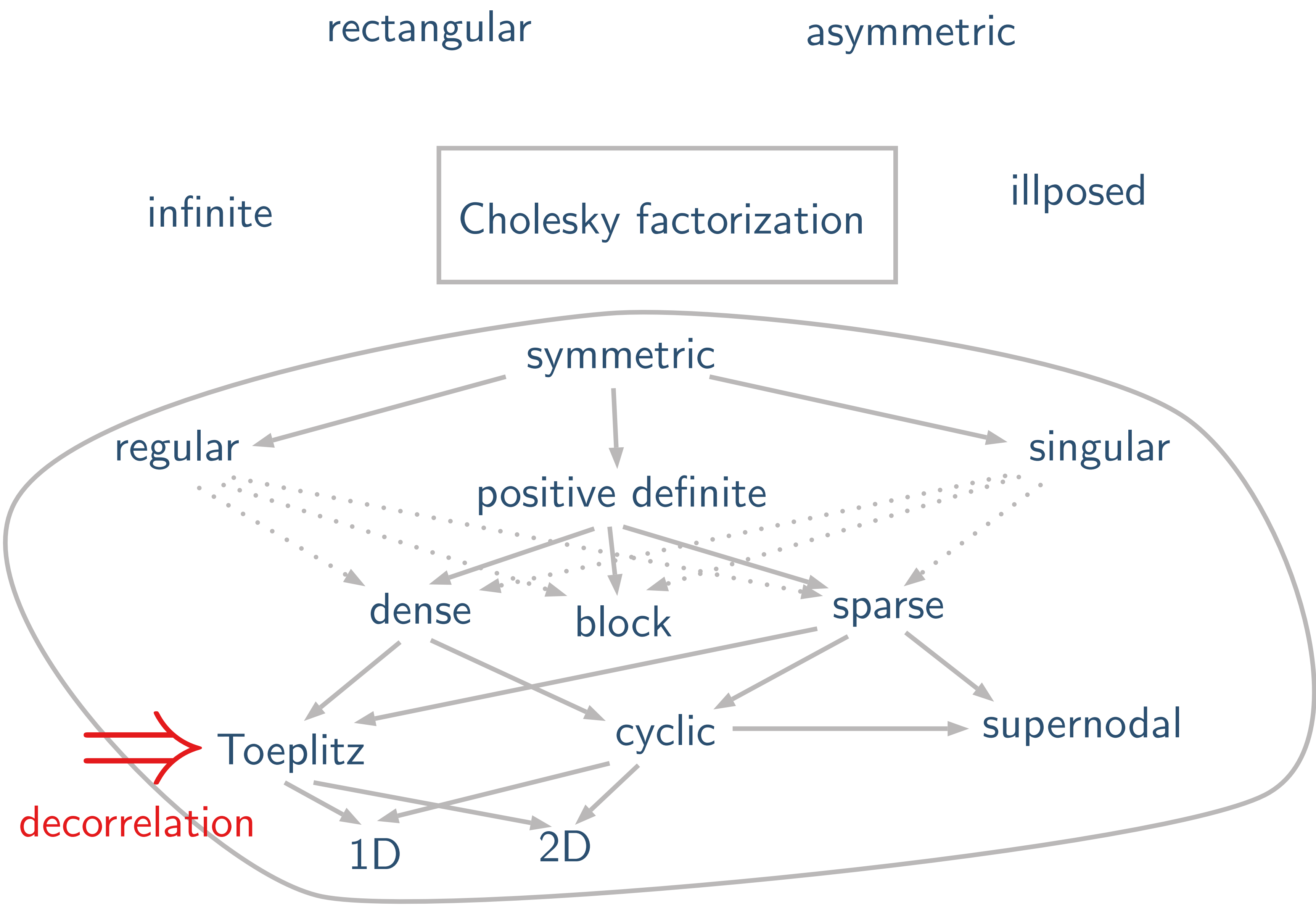
Cholesky application



Cholesky application



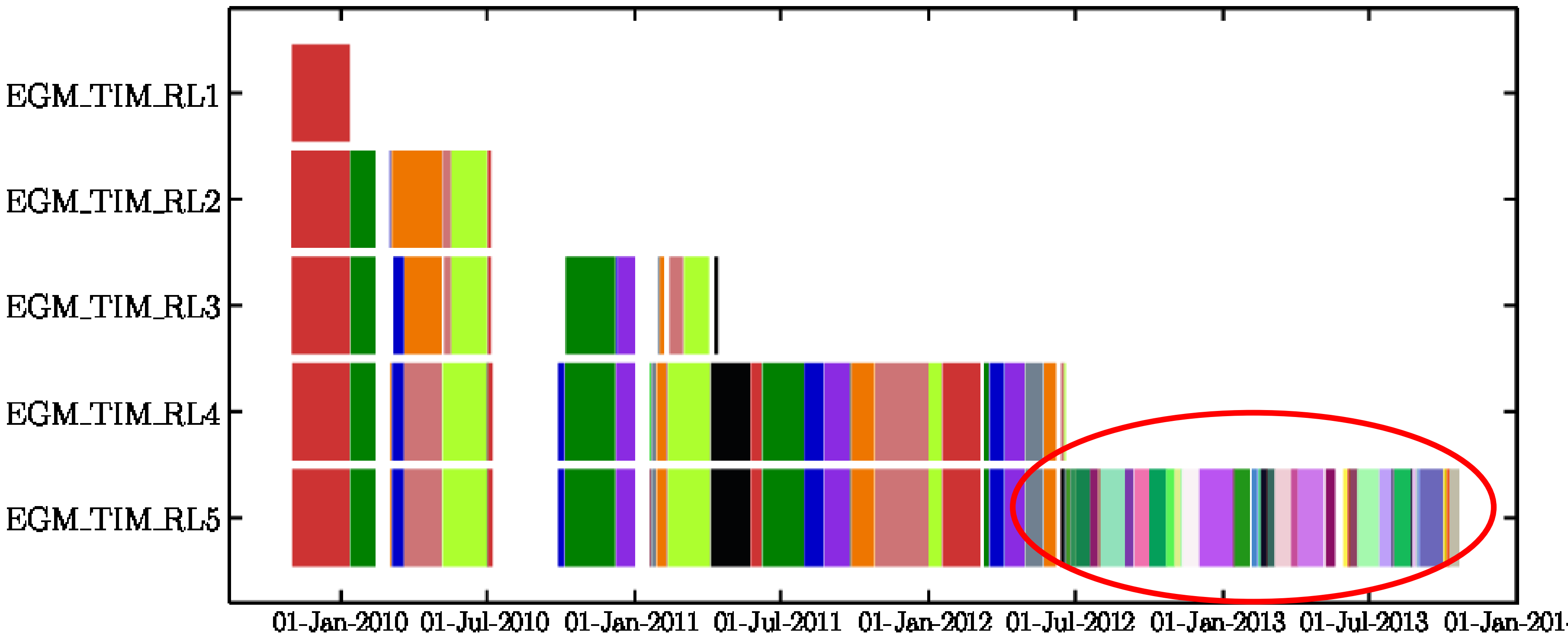
Cholesky application



Example: GOCE processing

GOCE-TIM data segments

440 mio observations, eff. 1273 days, 88 segments



Special focus on LOOC (13-Jun-2012 — 21-Oct-2013)

4 × 40,578,861 observations, eff. 463 days, 48 segments

LOOC-only models have the same accuracy level as the TIM_RL4 model

Gauss-Markov model — complete decorrelation

input:	$\boldsymbol{\ell},\ \boldsymbol{\Sigma}$	$\bar{\boldsymbol{\ell}} = \boldsymbol{H}\boldsymbol{\ell},\ \bar{\boldsymbol{\Sigma}} = \boldsymbol{I}$
model:	$\boldsymbol{\ell} + \boldsymbol{v} = \boldsymbol{A}\,\boldsymbol{x}$	$\bar{\boldsymbol{\ell}} + \bar{\boldsymbol{v}} = \bar{\boldsymbol{A}}\boldsymbol{x}$
principle:	$\boldsymbol{v}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{v}$	$\bar{\boldsymbol{v}}^T\bar{\boldsymbol{v}}$

$$\left(\boldsymbol{R}^{-1}\right)^T = \boldsymbol{H}$$

$$\boldsymbol{\Sigma} = \boldsymbol{R}^T\boldsymbol{R}$$

$$\bar{\boldsymbol{\ell}} = \left(\boldsymbol{R}^{-1}\right)^T\boldsymbol{\ell}$$

$$\bar{\boldsymbol{A}} = \left(\boldsymbol{R}^{-1}\right)^T\boldsymbol{A}$$

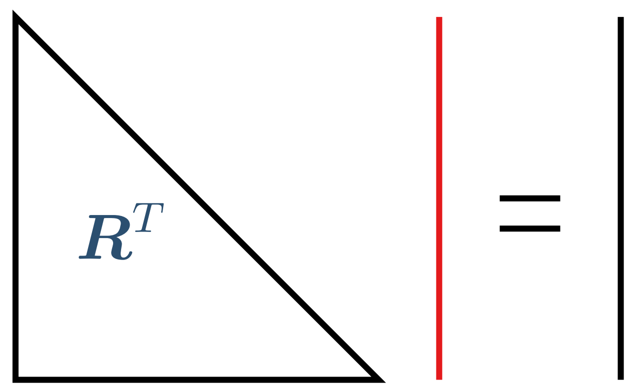
$$\bar{\boldsymbol{\Sigma}} = \left(\boldsymbol{R}^{-1}\right)^T\boldsymbol{\Sigma}\boldsymbol{R}^{-1} = \boldsymbol{I}$$

Gauss-Markov model — complete decorrelation

$$\bar{\boldsymbol{\ell}} = \boldsymbol{H}\boldsymbol{\ell}$$

$$\bar{\boldsymbol{\ell}} = (\boldsymbol{R}^{-1})^T\boldsymbol{\ell}$$

$$\boldsymbol{R}^T\bar{\boldsymbol{\ell}} = \boldsymbol{\ell}$$



$$\bar{\ell}_t = \frac{1}{r_{tt}}\left(\ell_t - \sum_{i=1}^{t-1} r_{it}\bar{\ell}_i\right)$$

$\Longrightarrow$ **linear - time variant - recursive - causal - filter**

Gauss-Markov model — complete decorrelation

pros and cons: recursive decorrelation filters

+ finite filter — finite time series
+ flexible with respect to data distribution (data gaps)
+ utilization of sparse structures — finite covariance functions

− recursive procedures are time-consuming
− time variant filter - attended memory $\mathcal{O}\{n^2\}$
− computational complexity $\mathcal{O}\{n^3\}$ $\frac{1}{3}(n^3 + 8n)$
− parallel implementation is very elaborately
− scalability $\mathcal{O}\{\sqrt{cores}\}$

Decorrelation — $\boldsymbol{\Sigma} \Longrightarrow \boldsymbol{H}$

crucial issue: numerical implementation of the factorization $\boldsymbol{\Sigma} \Longrightarrow \boldsymbol{H}$

Decomposition of	Computation of $\boldsymbol{R}$ or $\bar{\boldsymbol{R}}$	Decorrelation matrix	Decorrelation of $\boldsymbol{\ell}$
$\boldsymbol{\Sigma} = \boldsymbol{R}^T\boldsymbol{R}$	Forward Cholesky reduction	$\boldsymbol{H} = (\boldsymbol{R}^T)^{-1}$	$\bar{\boldsymbol{\mathcal{L}}} = (\boldsymbol{R}^T)^{-1}\boldsymbol{\mathcal{L}} \Longleftrightarrow \boldsymbol{R}^T\bar{\boldsymbol{\mathcal{L}}} = \boldsymbol{\mathcal{L}}$ $\begin{bmatrix} * & & \\ * & \boldsymbol{R}^T & \\ * & * & * \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{\mathcal{L}}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{L}} \end{bmatrix}$ Causal recursive filter
$\boldsymbol{\Sigma} = \bar{\boldsymbol{R}}\bar{\boldsymbol{R}}^T$	Backward Cholesky reduction	$\boldsymbol{H} = \bar{\boldsymbol{R}}^{-1}$	$\bar{\boldsymbol{\mathcal{L}}} = \bar{\boldsymbol{R}}^{-1}\boldsymbol{\mathcal{L}} \Longleftrightarrow \bar{\boldsymbol{R}}\bar{\boldsymbol{\mathcal{L}}} = \boldsymbol{\mathcal{L}}$ $\begin{bmatrix} * & * & * \\ & \bar{\boldsymbol{R}} & * \\ & & * \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{\mathcal{L}}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{L}} \end{bmatrix}$ Anti-causal recursive filter
$\boldsymbol{\Sigma} = (\boldsymbol{R}^{-1})^T\boldsymbol{R}^{-1}$	Recursive backward edging	$\boldsymbol{H} = \boldsymbol{R}$	$\bar{\boldsymbol{\mathcal{L}}} = \boldsymbol{R}\boldsymbol{\mathcal{L}}$ $\begin{bmatrix} \bar{\boldsymbol{\mathcal{L}}} \end{bmatrix} = \begin{bmatrix} * & * & * \\ & \boldsymbol{R} & * \\ & & * \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{L}} \end{bmatrix}$ Anti-causal non-recursive filter
$\boldsymbol{\Sigma} = \bar{\boldsymbol{R}}^{-1}(\bar{\boldsymbol{R}}^{-1})^T$	Recursive forward edging	$\boldsymbol{H} = \bar{\boldsymbol{R}}^T$	$\bar{\boldsymbol{\mathcal{L}}} = \bar{\boldsymbol{R}}^T\boldsymbol{\mathcal{L}}$ $\begin{bmatrix} \bar{\boldsymbol{\mathcal{L}}} \end{bmatrix} = \begin{bmatrix} * & & \\ * & \boldsymbol{R}^T & \\ * & * & * \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{L}} \end{bmatrix}$ Causal non-recursive filter

Toeplitz systems — Cholesky reduction

Focus: Causal filter via Cholesky reduction of Toeplitz structured covariance matrices resulting from AR(p) process:

$$\mathcal{S}_t = \alpha_1 \mathcal{S}_{t-1} + \alpha_2 \mathcal{S}_{t-2} + \ldots + \alpha_p \mathcal{S}_{t-p} + \mathcal{E}_t; \quad \boldsymbol{\mathcal{E}} \sim WN(\mathbf{0}, \boldsymbol{I}\sigma^2_{\mathcal{E}})$$

Toeplitz systems — Cholesky reduction

Focus: Causal filter via Cholesky reduction of Toeplitz structured covariance matrices resulting from AR(p) process:

$$\mathcal{S}_t = \alpha_1 \mathcal{S}_{t-1} + \alpha_2 \mathcal{S}_{t-2} + \ldots + \alpha_p \mathcal{S}_{t-p} + \mathcal{E}_t; \quad \boldsymbol{\mathcal{E}} \sim WN(\mathbf{0}, \boldsymbol{I}\sigma^2_{\mathcal{E}})$$

Yule-Walker equation

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_p & \cdots & \gamma_n \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-1} & \cdots & \gamma_{n-1} \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ \gamma_p & \gamma_{p-1} & \cdots & \gamma_0 & \cdots & \gamma_{p-1} \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ \gamma_n & \gamma_{n-1} & \cdots & \gamma_{p-1} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\alpha_1 \\ \vdots \\ -\alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma^2_{\mathcal{E}} \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Focus: Causal filter via Cholesky reduction of Toeplitz structured covariance matrices resulting from AR(p) process:

$$\mathcal{S}_t = \alpha_1 \mathcal{S}_{t-1} + \alpha_2 \mathcal{S}_{t-2} + \ldots + \alpha_p \mathcal{S}_{t-p} + \mathcal{E}_t; \quad \boldsymbol{\mathcal{E}} \sim WN(\mathbf{0}, \boldsymbol{I}\sigma^2_{\mathcal{E}})$$

Yule-Walker equation

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_p & \cdots & \gamma_n \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-1} & \cdots & \gamma_{n-1} \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ \gamma_p & \gamma_{p-1} & \cdots & \gamma_0 & \cdots & \gamma_{p-1} \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ \gamma_n & \gamma_{n-1} & \cdots & \gamma_{p-1} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\alpha_1 \\ \vdots \\ -\alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma^2_{\mathcal{E}} \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Causal non-recursive filter

(written in reversed order)

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_j \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{j-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_j & \gamma_{j-1} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} \bar{r}^{(-1)}_{jj} \bar{r}^{(-1)}_{jj} \\ \bar{r}^{(-1)}_{j-1,j} \bar{r}^{(-1)}_{jj} \\ \vdots \\ \bar{r}^{(-1)}_{1j} \bar{r}^{(-1)}_{jj} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Toeplitz systems — Cholesky reduction

Focus: Causal filter via Cholesky reduction of Toeplitz structured covariance matrices resulting from AR(p) process:

$$\mathcal{S}_t = \alpha_1 \mathcal{S}_{t-1} + \alpha_2 \mathcal{S}_{t-2} + \ldots + \alpha_p \mathcal{S}_{t-p} + \mathcal{E}_t; \quad \boldsymbol{\mathcal{E}} \sim WN(\mathbf{0}, \boldsymbol{I}\sigma_{\mathcal{E}}^2)$$

Yule-Walker equation

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_p & \cdots & \gamma_n \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-1} & \cdots & \gamma_{n-1} \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ \gamma_p & \gamma_{p-1} & \cdots & \gamma_0 & \cdots & \gamma_{p-1} \\ \vdots & \vdots & \ddots & & \ddots & \vdots \\ \gamma_n & \gamma_{n-1} & \cdots & \gamma_{p-1} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\alpha_1 \\ \vdots \\ -\alpha_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_{\mathcal{E}}^2 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Causal non-recursive filter

(written in reversed order)

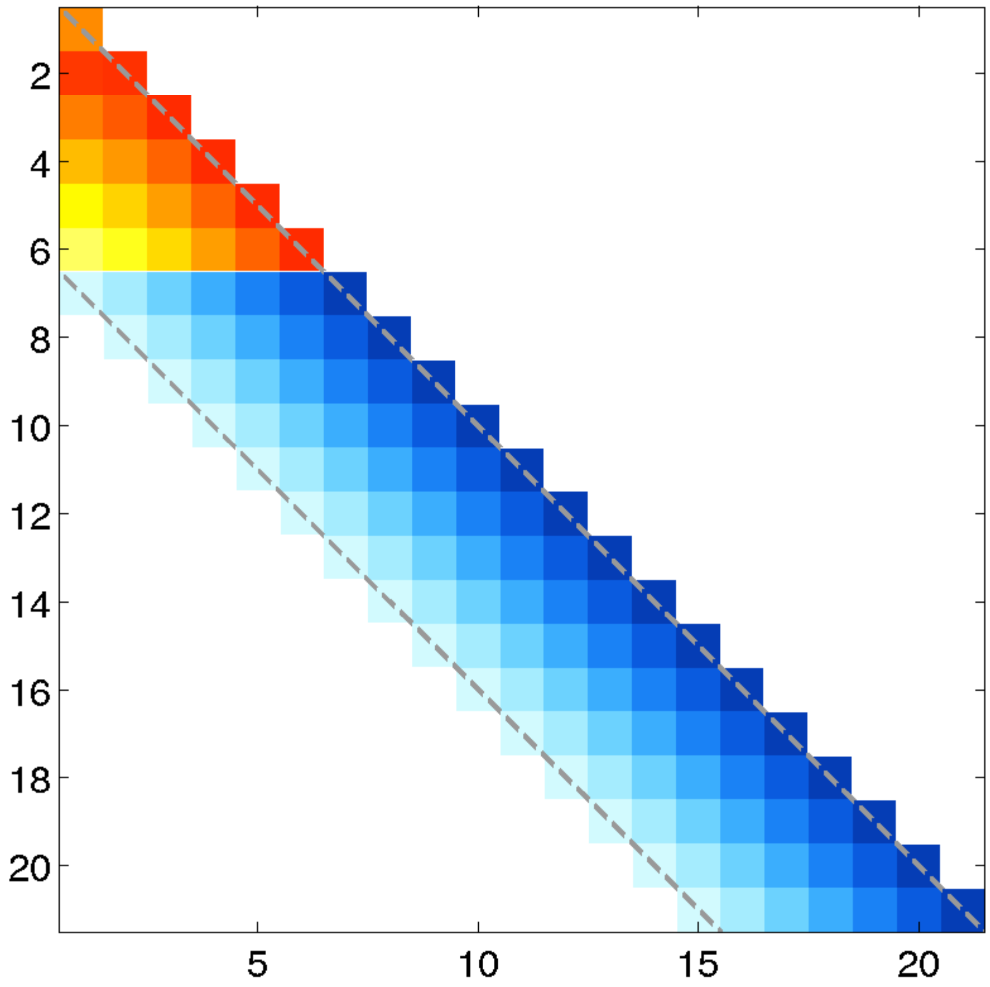
$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_j \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{j-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_j & \gamma_{j-1} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} \bar{r}_{jj}^{(-1)} \bar{r}_{jj}^{(-1)} \\ \bar{r}_{j-1,j}^{(-1)} \bar{r}_{jj}^{(-1)} \\ \vdots \\ \bar{r}_{1j}^{(-1)} \bar{r}_{jj}^{(-1)} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Resumé: Efficient computation of the filter coefficients

$j < p : \bar{r}_{ij}^{(-1)},\ i=1,\ldots,j$ derived from recursive forward edging (left box)

$j \geq p : \bar{r}_{jj}^{(-1)} = \frac{1}{\sigma_{\mathcal{E}}};\quad \bar{r}_{j-k,j}^{(-1)} = -\frac{\alpha_k}{\sigma_{\mathcal{E}}},\ k=1,\ldots,p$ fixed by coefficient comparison

Causal non-recursive filters: Rigorous warmup strategy for **finite** time series



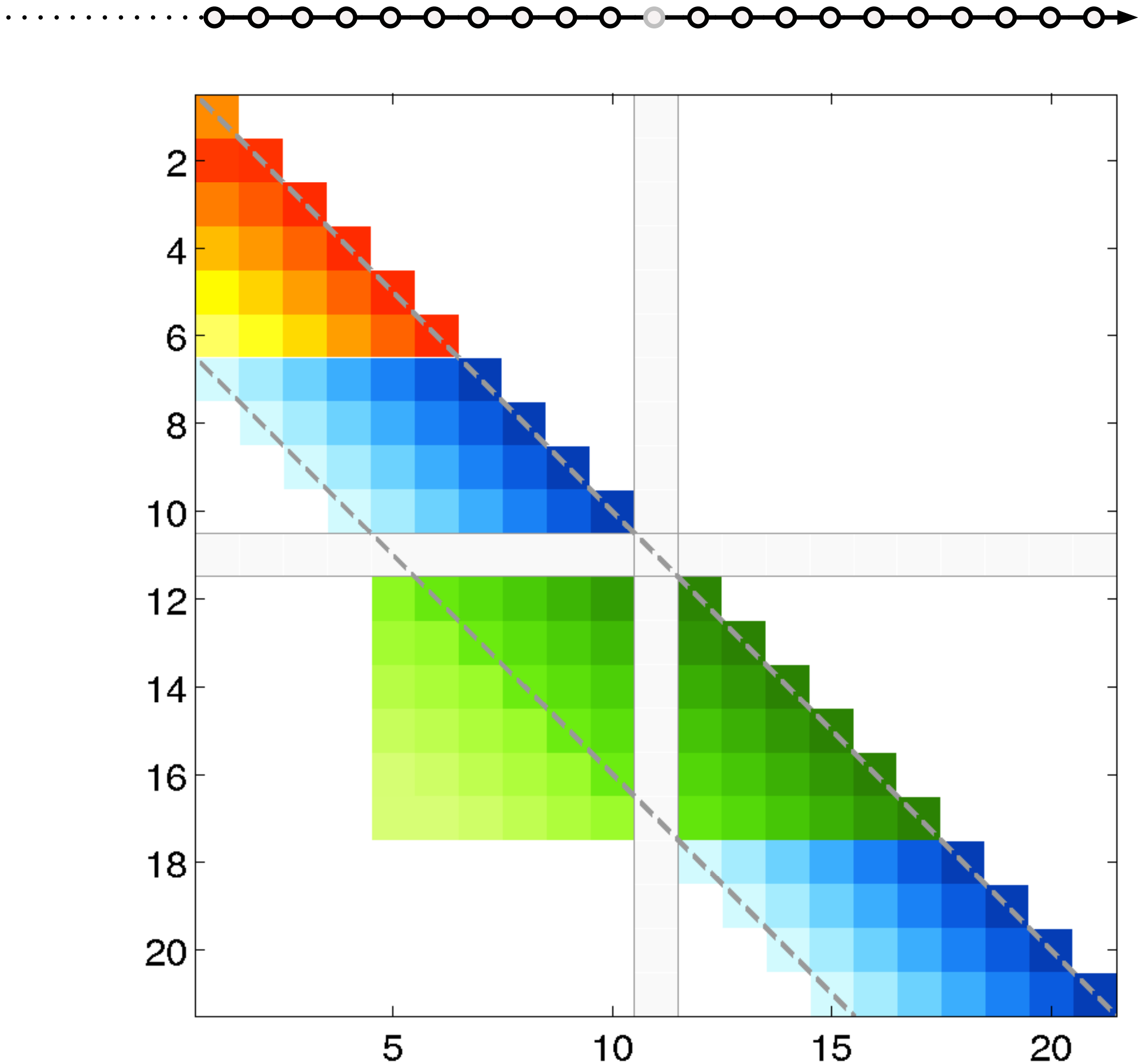
sparse filter with almost Toeplitz structure

warmup:
computed by Cholesky recursive forward edging

main part:
results direct from the AR coefficients

Adaptive whitening filter

Adaptive causal non-recursive whitening filter: warmup and single gap



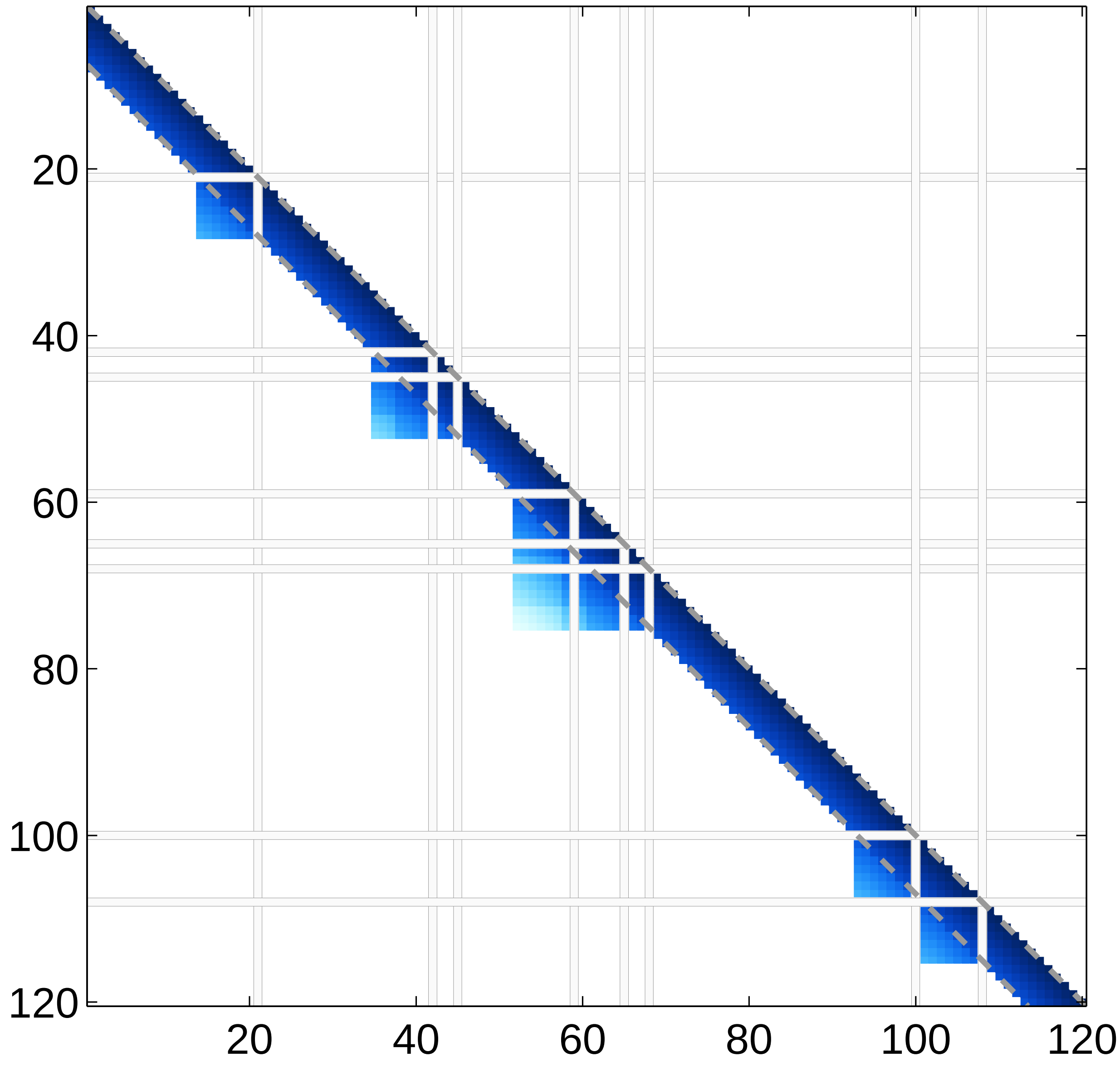
sparse filter with almost Toeplitz structure

main part:
results direct from the AR coefficients

warmup:
computed by Cholesky recursive forward edging

data gaps:
computed by Cholesky recursive forward edging

Adaptive causal non-recursive whitening filter: multiple gaps



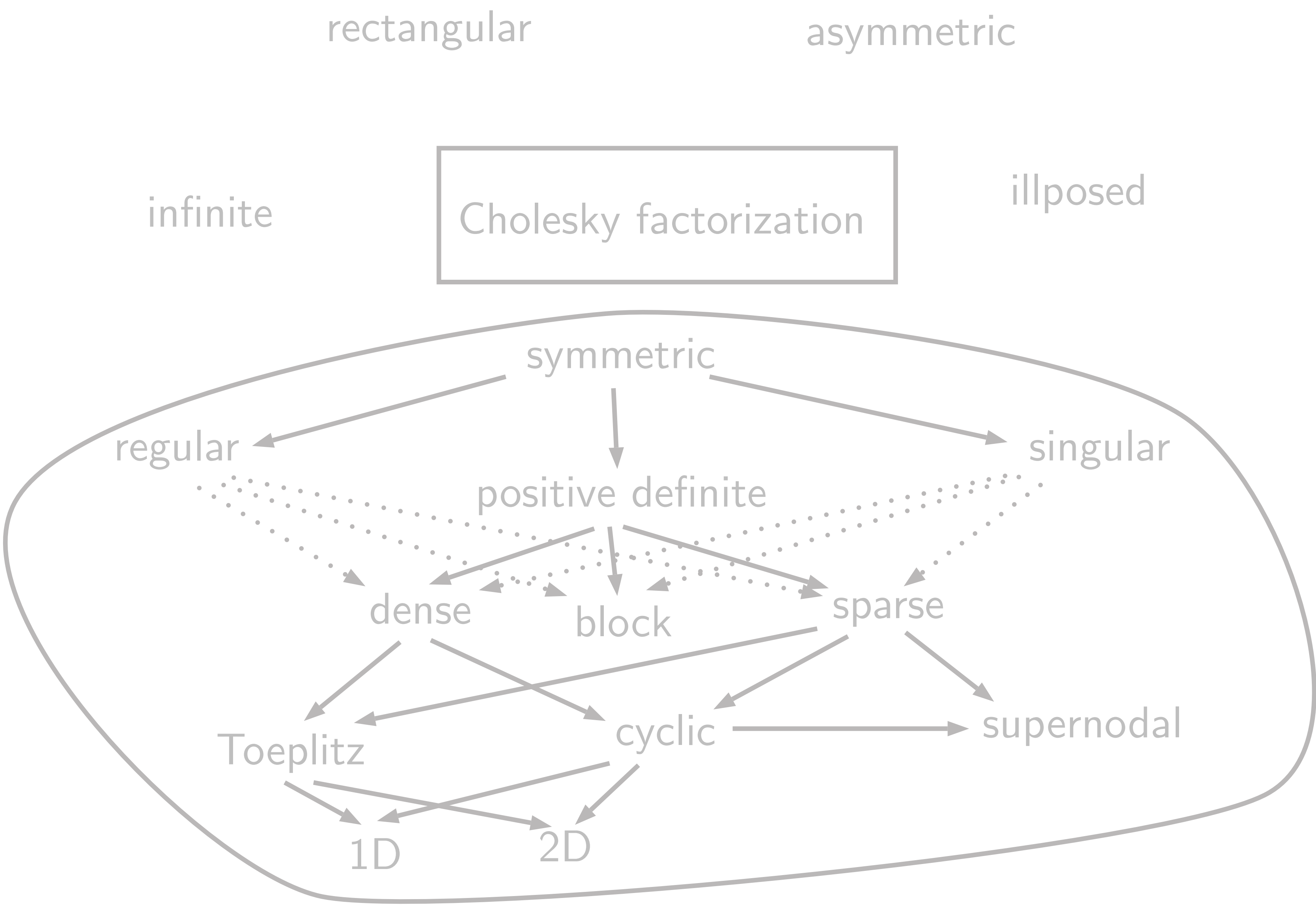
sparse filter with almost Toeplitz structure

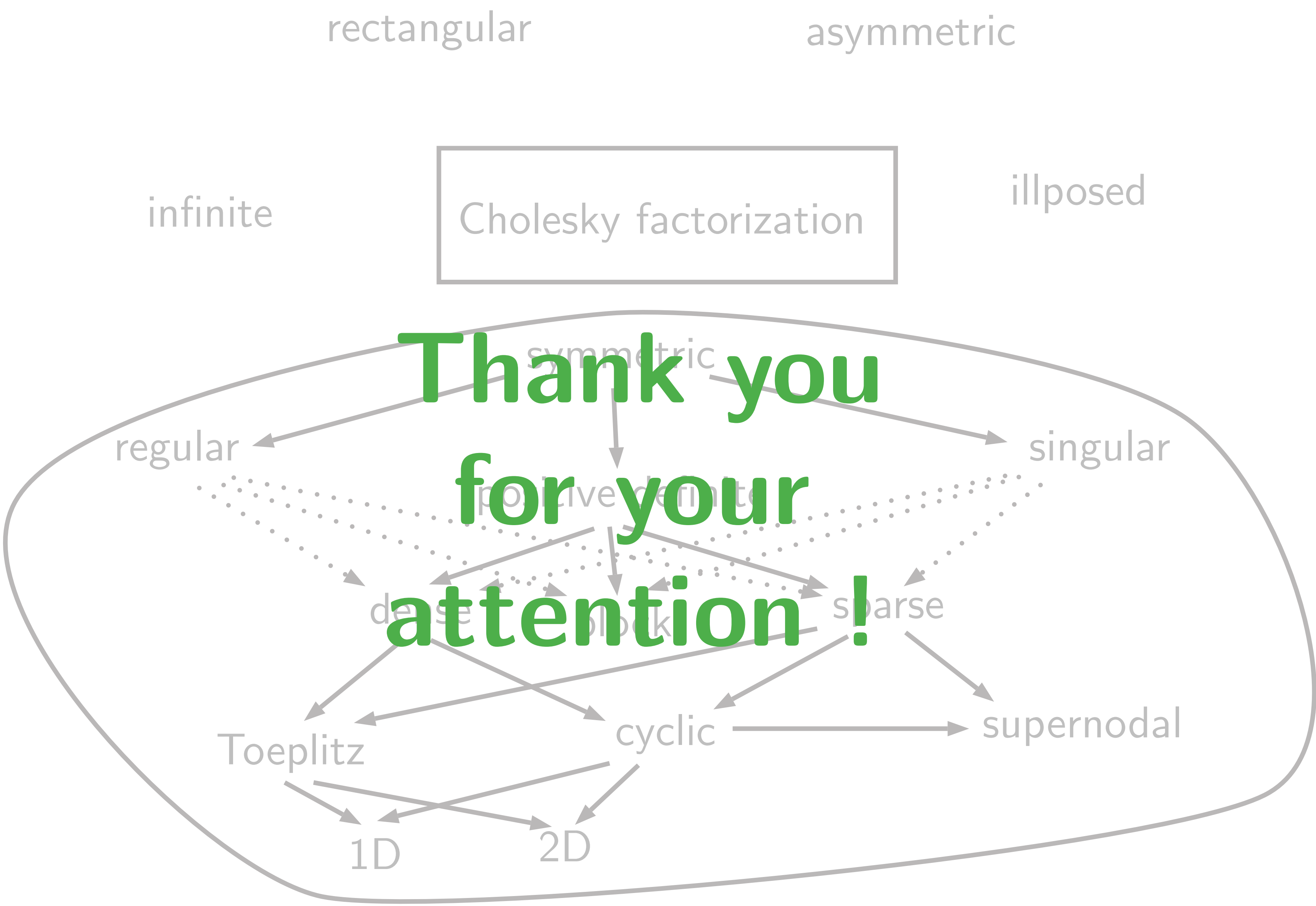
characteristics:

#data points : 120
#model : AR(7)-process
#gaps : 21
42 45
59 65 68
100
108

Resumé

Resumé


rectangular
asymmetric
infinite
Cholesky factorization
illposed
symmetric
regular
singular
positive definite
dense
block
sparse
Toeplitz
cyclic
supernodal
1D
2D



Thank you for your attention !