Gaussian processes, part 1: multivariate Gaussians

Stephen Huan

https://cgdct.moe

Theory club 2023-09-22

Overview

Cholesky factorization

Characterizations

References

Symmetric LU decomposition

$M = LU$ where L is lower triangular and U is upper triangular

Not always possible, need PLU in general!

Special case for symmetric matrices

Theorem

If $M = M^\top$ and $\det(M) \neq 0$, then $M = LDL^\top$ where L is from the LU decomposition of M and D is the diagonal of U.

Proof sketch.

(MATH3406 Fall 2021, Prof. Wing Li) Let $M = LDK$. Just do matrix multiplication on $M = M^\top \Rightarrow (LDK) = (LDK)^\top$. From matrix multiplication, able to see $K = L^\top$. □

Cholesky factorization

Let M be (symmetric) *positive definite*, that is,

$$\langle \boldsymbol{x}, M\boldsymbol{x} \rangle = \langle M\boldsymbol{x}, \boldsymbol{x} \rangle \geq 0 \text{ for all } \boldsymbol{x} \in \mathbb{R}^n.$$

Then $M = LDL^\top$ becomes $LL^\top$,

$$M = LDL^\top = L(\sqrt{D}\sqrt{D})L^\top = (L\sqrt{D})(L\sqrt{D})^\top = L'L'^\top$$

for $L' := L\sqrt{D}$ and $(\sqrt{D})_{i,j} := \sqrt{D_{i,j}}$.

This is the Cholesky factorization!

Why should we care?

$\Theta = LL^\top$, L has N columns, s non-zero entries per column

$L\boldsymbol{v}$ and $L^{-1}\boldsymbol{v}$ both cost $\mathcal{O}(Ns)$ floating-point operations

- Matrix-vector product $\Theta\boldsymbol{v} \to L(L^\top \boldsymbol{v})$
  - $N^2 \to Ns$
- Solving linear system $\Theta^{-1}\boldsymbol{v} \to L^{-\top}(L^{-1}\boldsymbol{v})$
  - $N^3 \to Ns$
- Log determinant $\operatorname{logdet}\Theta \to 2\operatorname{logdet} L = 2\sum_{i=1}^{N} \log L_{i,i}$
  - $N^3 \to N$
- Sampling from $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Theta) \to \boldsymbol{z} \sim \mathcal{N}(\mathbf{0}, \mathrm{Id}), \boldsymbol{x} = L\boldsymbol{z} + \boldsymbol{\mu}$
  - ??? $\to Ns$

Here, we care about its statistical interpretation

Schur complement

Block $\Theta = \begin{pmatrix} \Theta_{1,1} & \Theta_{1,2} \\ \Theta_{2,1} & \Theta_{2,2} \end{pmatrix}$ then perform a step of Gaussian elimination

$$\begin{pmatrix} \Theta_{1,1} & \Theta_{1,2} \\ \mathbf{0} & \Theta_{2,2} - \Theta_{2,1}\Theta_{1,1}^{-1}\Theta_{1,2} \end{pmatrix}.$$

Denote the term in blue, the *Schur complement* of Θ on $\Theta_{1,1}$, as

$$\Theta = \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{2,1}\Theta_{1,1}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{1,1} & 0 \\ 0 & \Theta_{2,2|1} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Theta_{1,1}^{-1}\Theta_{1,2} \\ 0 & \mathrm{Id} \end{pmatrix},$$

so we see the Cholesky factorization of Θ is

$$\mathrm{chol}(\Theta) = \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{2,1}\Theta_{1,1}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \mathrm{chol}(\Theta_{1,1}) & 0 \\ 0 & \mathrm{chol}(\Theta_{2,2|1}) \end{pmatrix}.$$

Recursing on both diagonal blocks finishes the construction.

Determinant of a block matrix

$$\Theta = \begin{pmatrix} \Theta_{1,1} & \Theta_{1,2} \\ \Theta_{2,1} & \Theta_{2,2} \end{pmatrix}$$

$$\det(\Theta) = \det(\Theta_{1,1})\det(\Theta_{2,2}) - \det(\Theta_{2,1})\det(\Theta_{1,2})? \qquad \text{wrong!}$$

$$= \det(\Theta_{1,1}\Theta_{2,2} - \Theta_{2,1}\Theta_{1,2})? \qquad \text{wrong!}$$

Schur complement gives proper answer

$$\Theta = \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{2,1}\Theta_{1,1}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{1,1} & 0 \\ 0 & \Theta_{2,2|1} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Theta_{1,1}^{-1}\Theta_{1,2} \\ 0 & \mathrm{Id} \end{pmatrix}$$

$$\det(\Theta) = \det(\Theta_{1,1})\det\left(\Theta_{2,2|1}\right)$$

Submatrix of inverse

$$\Theta = \begin{pmatrix} \Theta_{1,1} & \Theta_{1,2} \\ \Theta_{2,1} & \Theta_{2,2} \end{pmatrix}$$

$$(\Theta^{-1})_{2,2} = ?$$

$$= (\Theta_{2,2})^{-1}? \qquad \text{wrong!}$$

Schur complement to the rescue again!

Proper Submatrix of Inverse

Recall the factorization of Θ as

$$\begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{2,1}\Theta_{1,1}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{1,1} & 0 \\ 0 & \Theta_{2,2|1} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Theta_{1,1}^{-1}\Theta_{1,2} \\ 0 & \mathrm{Id} \end{pmatrix}.$$

Inverting both sides of the equation, Θ^{-1} is

$$\begin{pmatrix} \mathrm{Id} & -\Theta_{1,1}^{-1}\Theta_{1,2} \\ 0 & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{1,1}^{-1} & 0 \\ 0 & \Theta_{2,2|1}^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & 0 \\ -\Theta_{2,1}\Theta_{1,1}^{-1} & \mathrm{Id} \end{pmatrix}$$

$$= \begin{pmatrix} \Theta_{1,1}^{-1} + \left(\Theta_{1,1}^{-1}\Theta_{1,2}\right)\Theta_{2,2|1}^{-1}\left(\Theta_{2,1}\Theta_{1,1}^{-1}\right) & -\left(\Theta_{1,1}^{-1}\Theta_{1,2}\right)\Theta_{2,2|1}^{-1} \\ -\Theta_{2,2|1}^{-1}\left(\Theta_{2,1}\Theta_{1,1}^{-1}\right) & \Theta_{2,2|1}^{-1} \end{pmatrix}$$

So $(\Theta^{-1})_{2,2}$ can be read off as $\Theta_{2,2|1}^{-1} := (\Theta_{2,2|1})^{-1}$.

Leaving linear algebra...

Is the Schur complement symmetric positive definite (s.p.d.)?

- Have been assuming so throughout

Is Schur complementing transitive?

- i.e. suppose we have Θ blocked as

$$\Theta = \begin{pmatrix} \Theta_{1,1} & \Theta_{1,2} & \Theta_{1,3} \\ \Theta_{2,1} & \Theta_{2,2} & \Theta_{2,3} \\ \Theta_{3,1} & \Theta_{3,2} & \Theta_{3,3} \end{pmatrix}$$

- Is Θ complemented on $\Theta_{1,1}$ and then on $\Theta_{2,2}$ the same as Θ complemented on $\begin{pmatrix} \Theta_{1,1} & \Theta_{1,2} \\ \Theta_{2,1} & \Theta_{2,2} \end{pmatrix}$? Intuitively yes, but tedious

Intuitive and economic proofs from statistical perspective

Multivariate Gaussian

Recall: Gaussian (or normal) distribution:

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

$$\pi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

(alternative) Defining property: completely determined by mean and variance, all higher-order cumulants zero.

We're going to extend this to higher dimensions. Consider

$$\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$$

where $\boldsymbol{x}$ ("variables") is a $N \times 1$ vector, $\boldsymbol{\mu}$ ("mean vector") is a $N \times 1$ vector, and Σ ("covariance matrix") is a $N \times N$ matrix

Defining everything

Naturally,

$$
\begin{aligned}
\mu_i &\coloneqq \mathbb{E}[x_i] \\
\boldsymbol{\mu} &= \mathbb{E}[\boldsymbol{x}] \\
\Sigma_{i,j} &\coloneqq \mathbb{C}\mathrm{ov}[x_i, x_j] \\
&= \mathbb{E}[(x_i - \mathbb{E}[x_i])(x_j - \mathbb{E}[x_j])] \\
&= \mathbb{E}[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top]
\end{aligned}
$$

Two natural questions from here.

1. What is the probability density function (PDF) $\pi(\boldsymbol{x})$?
2. How can we sample from $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$?

A technical path

Recall *moment/cumulant generating functions*

$$M_{\boldsymbol{X}}(\boldsymbol{\xi}) \coloneqq \mathbb{E}[\exp\Big(\boldsymbol{\xi}^\top \boldsymbol{x}\Big)]$$
$$K_{\boldsymbol{X}}(\boldsymbol{\xi}) \coloneqq \log M_{\boldsymbol{X}}(\boldsymbol{\xi})$$

One characterization: multivariate Gaussian has

$$M_{\boldsymbol{x}}(\boldsymbol{\xi}) = \exp\Big(\xi^\top \boldsymbol{\mu} + \boldsymbol{\xi}^\top \Theta^{-1} \boldsymbol{\xi}\Big)$$
$$K_{\boldsymbol{x}}(\boldsymbol{\xi}) = \xi^\top \boldsymbol{\mu} + \boldsymbol{\xi}^\top \Theta^{-1} \boldsymbol{\xi}$$

Invert the integral transformation to recover PDF

Instead: exploit moment matching and independence

Identity covariance

If Σ has the simplest structure possible, it's trivial:

$$\boldsymbol{z} \sim \mathcal{N}(\mathbf{0}, \mathrm{Id}_N)$$

$$z_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

$$\begin{aligned} f(\boldsymbol{z}) &= \prod_{i=1}^{N} f(z_i) \\ &= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z_i^2} \\ &= \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{1}{2}(z_1^2 + \cdots + z_N^2)} \\ &= \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{z}} \end{aligned}$$

Moment matching

Generalize to arbitrary Σ by moment matching

$$
\begin{aligned}
\boldsymbol{z} &\sim \mathcal{N}(\mathbf{0}, \mathrm{Id}_N) \\
\boldsymbol{x} &= L\boldsymbol{z} + \boldsymbol{\mu} \\
\mathbb{E}[\boldsymbol{x}] &= \mathbb{E}[L\boldsymbol{z} + \boldsymbol{\mu}] = L\,\mathbb{E}[\boldsymbol{z}] + \boldsymbol{\mu} = \boldsymbol{\mu} \\
\mathbb{C}\mathrm{ov}[\boldsymbol{x}] &= \mathbb{E}[(\boldsymbol{x} - \mathbb{E}[\boldsymbol{x}])(\boldsymbol{x} - \mathbb{E}[\boldsymbol{x}])^\top] \\
&= \mathbb{E}[L\boldsymbol{z}(L\boldsymbol{z})^\top] \\
&= \mathbb{E}[L\boldsymbol{z}\boldsymbol{z}^\top L^\top] \\
&= L\,\mathbb{E}[\boldsymbol{z}\boldsymbol{z}^\top]L^\top \\
&= LL^\top
\end{aligned}
$$

so $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, LL^\top)$. We want $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, so $\Sigma = LL^\top$.

Sampling with Cholesky Factorization

As we just saw, we can sample $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ by instead sampling $\boldsymbol{z} \sim \mathcal{N}(\mathbf{0}, I_N)$ and computing $\boldsymbol{x} = L\boldsymbol{z} + \boldsymbol{\mu}$.

Since $LL^\top = \Sigma$, a natural pick is $L = \mathrm{chol}(\Sigma)$.

Why is Σ s.p.d.? Because it's a covariance/Gram matrix.

$$
\begin{aligned}
\Sigma &= \mathbb{E}[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top] \\
\boldsymbol{y}^\top \Sigma \boldsymbol{y} &= \boldsymbol{y}^\top \mathbb{E}[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top] \boldsymbol{y} \\
&= \mathbb{E}[\boldsymbol{y}^\top (\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{y}] \\
&= \mathbb{E}[((\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{y})^\top (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{y}] \\
&= \mathbb{E}[\|(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{y}\|^2] \geq 0
\end{aligned}
$$

Pushforward

Integrate against test function ψ

$$\mathbb{E}_\eta[\psi \circ T] \coloneqq \int [\psi \circ T](\boldsymbol{z})\, \eta(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} = \int \psi(\boldsymbol{x})\, T_\sharp \eta(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$

with change of variables $\boldsymbol{x} = T(\boldsymbol{z})$ and pushforward

$$T_\sharp \eta \coloneqq (\eta \circ T^{-1}) \left| \det \nabla T^{-1} \right|$$

Therefore we have

$$\mathbb{E}_\eta[\psi \circ T] = \int \psi(\boldsymbol{x})\, T_\sharp \eta(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int \psi(\boldsymbol{x})\, \pi(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \mathbb{E}_\pi[\psi]$$

if and only if $T_\sharp \eta = \pi$.

PDF from transport

$$\pi(\boldsymbol{z}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{z}}$$

Expanding $\det\left(L^{-1}\right) f(L^{-1}(\boldsymbol{x} - \boldsymbol{\mu}))$,

$$= \frac{1}{\det(L)} f(L^{-1}(\boldsymbol{x} - \boldsymbol{\mu}))$$

$$= \frac{1}{\det(L)} \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{1}{2}(L^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))^\top (L^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))}$$

Since $LL^\top = \Sigma$, $\det(\Sigma) = \det(L)^2$

$$= \frac{1}{\sqrt{(2\pi)^N \det(\Sigma)}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top L^{-T} L^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}$$

$$= \frac{1}{\sqrt{(2\pi)^N \det(\Sigma)}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}$$

Summary

Compare PDFs of multivariate normal and scalar normal:

$$\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$$

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^N \det(\Sigma)}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

Compare to scalar:

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Remarkable similarity!

Cholesky factorization for Gaussians

Sampling: $\boldsymbol{x} = L\boldsymbol{z} + \mu$, matrix-vector product, $\mathcal{O}(Ns)$

(Log)-likelihood computation:

$$\begin{aligned}(\boldsymbol{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) &= (\boldsymbol{x} - \boldsymbol{\mu})^\top L^{-\top} L^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \\ &= (L^{-1}(\boldsymbol{x} - \boldsymbol{\mu}))^\top L^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \\ &= \|L^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\|^2\end{aligned}$$

Back-substitution, $\mathcal{O}(Ns)$

Closure of multivariate Gaussians

Many statistical operations preserve distribution

Affine transformation

Joint distribution & marginalization:

$$\boldsymbol{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_{1,1})$$
$$\boldsymbol{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \Sigma_{2,2})$$
$$\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix} \right)$$

Conditioning

Conditioning

Assume $\boldsymbol{\mu} = \mathbf{0}$ and use precision instead of covariance!

$$Q = \Sigma^{-1} = \begin{pmatrix} Q_{1,1} & Q_{1,2} \\ Q_{2,1} & Q_{2,2} \end{pmatrix}$$

$$\begin{aligned} \pi(\boldsymbol{x}_2 \mid \boldsymbol{x}_1) &= \frac{\pi(\boldsymbol{x}_1 \mid \boldsymbol{x}_2)\pi(\boldsymbol{x}_2)}{\pi(\boldsymbol{x}_1)} = \frac{\pi(\boldsymbol{x}_1, \boldsymbol{x}_2)}{\pi(\boldsymbol{x}_1)} \\ &\propto \pi(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ &\propto e^{-\frac{1}{2}\boldsymbol{x}_2^\top Q_{2,2}\boldsymbol{x}_2 - (Q_{2,1}\boldsymbol{x}_1)^\top \boldsymbol{x}_2} \end{aligned}$$

$$\boldsymbol{x}_2 \mid \boldsymbol{x}_1 \sim \mathcal{N}\left(-Q_{2,2}^{-1}Q_{2,1}\boldsymbol{x}_1, Q_{2,2}^{-1}\right)$$

If $\boldsymbol{\mu} \neq \mathbf{0}$, shift $\boldsymbol{x}^* = \boldsymbol{x} - \boldsymbol{\mu}$, $\mathbb{E}[\boldsymbol{x}^*] = \mathbf{0}$

$$\boldsymbol{x}_2 \mid \boldsymbol{x}_1 \sim \mathcal{N}\left(\boldsymbol{\mu}_2 - Q_{2,2}^{-1}Q_{2,1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1), Q_{2,2}^{-1}\right)$$

Conditioning with Schur Complements

$$
\begin{aligned}
\boldsymbol{x}_2 \mid \boldsymbol{x}_1 &\sim \mathcal{N}\left(\boldsymbol{\mu}_2 - Q_{22}^{-1} Q_{21}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1), Q_{22}^{-1}\right) \\
Q = \Sigma^{-1} &= \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \\
&= \begin{pmatrix} \Sigma_{11}^{-1} + \left(\Sigma_{11}^{-1}\Sigma_{12}\right)\Sigma_{22|1}^{-1}\left(\Sigma_{21}\Sigma_{11}^{-1}\right) & -\left(\Sigma_{11}^{-1}\Sigma_{12}\right)\Sigma_{22|1}^{-1} \\ -\Sigma_{22|1}^{-1}\left(\Sigma_{21}\Sigma_{11}^{-1}\right) & \Sigma_{22|1}^{-1} \end{pmatrix} \\
Q_{22}^{-1} &= (\Sigma_{22|1}^{-1})^{-1} = \Sigma_{22|1} \\
&= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \\
Q_{22}^{-1} Q_{21} &= -\Sigma_{22|1}(\Sigma_{22|1}^{-1}\Sigma_{21}\Sigma_{11}^{-1}) \\
&= -\Sigma_{21}\Sigma_{11}^{-1} \\
\boldsymbol{x}_2 \mid \boldsymbol{x}_1 &\sim \mathcal{N}\left(\boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)
\end{aligned}
$$

Statistical Interpretation

From conditioning,

$$\boldsymbol{x}_2 \mid \boldsymbol{x}_1 \sim \mathcal{N}\left(\boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$$

Schur complement $\iff$ conditional covariance!

s.p.d. because covariance matrices s.p.d.

Quotient rule statistically trivial:
$\pi((x_1 \mid x_2) \mid x_3) = \pi(x_1 \mid x_2, x_3)$

Conditioning in covariance $\iff$ marginalization in precision

References

- Bishop, Christopher M. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York: Springer. ISBN: 978-0-387-31073-2.
- Kallenberg, Olav (2021). *Foundations of Modern Probability*. Vol. 99. Probability Theory and Stochastic Modelling. Cham: Springer International Publishing. ISBN: 978-3-030-61870-4 978-3-030-61871-1. DOI: 10.1007/978-3-030-61871-1. (Visited on 09/15/2023).
- McCullagh, P (2018). *Tensor Methods in Statistics: Monographs on Statistics and Applied Probability*. Boca Raton, FL: CRC Press. ISBN: 978-1-351-08556-4 978-1-351-07711-8 978-1-351-09401-6 978-1-315-89801-8. (Visited on 04/17/2022).
- Rue, Havard and Leonhard Held (Feb. 2005). *Gaussian Markov Random Fields: Theory and Applications*. New York: Chapman and Hall/CRC. ISBN: 978-0-429-20882-9. DOI: 10.1201/9780203492024.