

Gaussian Processes: Cholesky Factorization and Causal Representation

Based on Antonio Sela's Lecture

April 4, 2025

Abstract

This document provides a detailed analysis of the covariance matrix of Gaussian processes, with particular emphasis on the Cholesky decomposition and its interpretation. We explore how the Cholesky factor of a positive definite covariance matrix provides a causal representation of a Gaussian process. This representation leads to an insightful connection between Gaussian processes and linear time-invariant systems, demonstrating how Gaussian processes can be viewed as the output of filtering white noise through a causal system.

1 Introduction to Gaussian Processes and Covariance Structure

A Gaussian process is a stochastic process where any finite collection of random variables has a multivariate normal distribution. The process is fully characterized by its mean function and covariance function. In this analysis, we focus on a Gaussian process with:

- A specified mean and standard deviation defining confidence intervals
- A covariance kernel in the form of an exponential multiplied by a trigonometric function
- Stationarity property (covariance depends only on the difference between points)

The covariance kernel has the form:

$$k(x_i, x_j) = \exp(-\alpha|x_i - x_j|) \cdot \cos(\omega|x_i - x_j|) \quad (1)$$

For discrete evaluation points, the covariance matrix \mathbf{K} is constructed by evaluating this kernel at all pairs of points. In our example, the process is discretized at 241 evenly spaced test points, resulting in a 241×241 covariance matrix.

2 Factorization of the Covariance Matrix

2.1 General Matrix Square Root Approach

The foundation of variance analysis is the factorization of the covariance matrix \mathbf{K} as:

$$\mathbf{K} = \mathbf{Q}\mathbf{Q}^T \quad (2)$$

With this factorization, we can introduce standard normal latent variables $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and define:

$$\mathbf{y} = \mathbf{Q}\mathbf{x} \quad (3)$$

The covariance matrix of \mathbf{y} is then:

$$\text{Cov}(\mathbf{y}) = \mathbb{E}[\mathbf{y}\mathbf{y}^T] \quad (4)$$

$$= \mathbb{E}[\mathbf{Q}\mathbf{x}\mathbf{x}^T\mathbf{Q}^T] \quad (5)$$

$$= \mathbf{Q}\mathbb{E}[\mathbf{x}\mathbf{x}^T]\mathbf{Q}^T \quad (6)$$

$$= \mathbf{Q}\mathbf{I}\mathbf{Q}^T \quad (7)$$

$$= \mathbf{Q}\mathbf{Q}^T = \mathbf{K} \quad (8)$$

This demonstrates that the latent variables \mathbf{x} explain the covariance structure observed in \mathbf{y} . It's important to note that matrix square roots are not unique (they differ by orthogonal transformations), so multiple explanations of the variance structure are possible.

2.2 Cholesky Decomposition

The Cholesky decomposition provides a specific factorization of a symmetric positive definite matrix. For our covariance matrix \mathbf{K} , the Cholesky decomposition gives a lower triangular matrix \mathbf{L} such that:

$$\mathbf{K} = \mathbf{L}\mathbf{L}^T \quad (9)$$

This can be computed in MATLAB using:

```
L = chol(K, 'lower');
```

We denote this lower triangular matrix as \mathbf{Q} in our subsequent analysis.

3 Causal Interpretation of the Cholesky Factor

3.1 Triangular Structure and Causality

The lower triangular structure of \mathbf{Q} enables a causal interpretation. When computing $\mathbf{y} = \mathbf{Q}\mathbf{x}$, each output y_i is influenced only by the latent variables x_1 through x_i :

$$y_i = \sum_{j=1}^i Q_{ij} x_j \quad (10)$$

Expanding this matrix product:

$$y_1 = Q_{11}x_1 \quad (11)$$

$$y_2 = Q_{21}x_1 + Q_{22}x_2 \quad (12)$$

$$y_3 = Q_{31}x_1 + Q_{32}x_2 + Q_{33}x_3 \quad (13)$$

$$\vdots \quad (14)$$

This structure has a causal interpretation:

- x_1 influences all outputs (y_1 through y_n)
- x_2 influences only y_2 through y_n (not y_1)
- x_3 influences only y_3 through y_n (not y_1 or y_2)

This causality means that each latent variable affects only the current and future observations, not past ones—a natural interpretation if we view the indices as time steps.

4 Analysis of the Cholesky Factor Structure

4.1 Column Structure

Each column of \mathbf{Q} represents the influence of a specific latent variable on all outputs. For stationary Gaussian processes, after the initial few columns (affected by boundary conditions), all subsequent columns exhibit the same pattern but shifted by one position.

For instance:

- Column 100 contains 99 zeros followed by coefficients $[h_0, h_1, h_2, \dots]$
- Column 101 contains 100 zeros followed by the same coefficients $[h_0, h_1, h_2, \dots]$
- Column 102 contains 101 zeros followed by the same coefficients $[h_0, h_1, h_2, \dots]$

This pattern indicates that the effect of each latent variable converges to a time-invariant structure, resembling the response of a linear time-invariant system.

4.2 Row Structure

Examining the rows of \mathbf{Q} provides a complementary perspective. For example, row 102 would have the structure:

$$[h_{102}, h_{101}, h_{100}, \dots, h_1, h_0, 0, 0, \dots, 0] \quad (15)$$

The zeros on the right reflect the causal property that future latent variables do not influence the current output. The non-zero coefficients show how past latent variables influence the current output, with their effect diminishing as we go further back in time (for stable systems).

5 Convolution Interpretation

5.1 Convolution Formula

The computation of y_i can be expressed as a convolution:

$$y_i = \sum_{s=0}^{i-1} h_s x_{i-s} \quad (16)$$

This is identical to the formula for calculating the output of a discrete-time linear time-invariant system with impulse response $\{h_s\}$ when the input is $\{x_i\}$.

In the context of our analysis, this means the Cholesky factor reveals that our Gaussian process can be represented as the output of a linear filter with impulse response derived from the Cholesky factor, applied to white noise (standard normal random variables).

5.2 Spectral Factorization Connection

For stationary processes, once boundary effects have vanished, the structure of the Cholesky factor reveals what is known as the spectral factor or causal square root of the process's spectrum. This establishes a direct connection between the Cholesky decomposition and spectral factorization in signal processing.

Formally, for a theoretically infinite process, the impulse response coefficients $\{h_s\}$ would extend infinitely into the past, but they decay to zero for stable systems, allowing practical approximation with finite memory.

6 Steady-State Behavior

In our analysis, after discarding the first handful of columns affected by boundary conditions, the remaining columns of the Cholesky factor exhibit a stationary pattern. This indicates that the system has reached a steady state where:

- All columns beyond a certain index have the same shape (just shifted)

- The impulse response derived from these columns represents the time-invariant filter that generates the Gaussian process from white noise
- This filter's characteristics directly relate to the spectral properties of the Gaussian process

For a stationary Gaussian process, this steady-state behavior is expected, as the covariance structure itself is translation-invariant.

7 Animation of the Causal Generation Process

The transcript describes an animation demonstrating the causal generation of a Gaussian process realization. This animation illustrates:

- The scaled rows of the \mathbf{Q} matrix, showing each latent variable's effect
- The properly scaled contribution of each latent variable to the overall process
- The progressive construction of a process realization from left to right
- The completed (fixed) part of the realization to the left of a moving boundary
- The currently evolving part to the right of this boundary

In this animation:

- Each new component has zero influence on past variables
- Components are added sequentially, building the realization causally
- The total variance is completely modeled to the left of a moving boundary
- Each new latent variable extends the completed portion of the realization

This visual representation reinforces the causal interpretation of the Cholesky factorization.

8 Practical Implementation in MATLAB

The analysis can be implemented in MATLAB with the following key steps:

1. Generate a covariance matrix using the specified kernel function:

```
K = zeros(n, n);
for i = 1:n
    for j = 1:n
        K(i,j) = exp(-alpha*abs(x(i)-x(j))) * cos(omega*abs(x(i)-x(j)));
    end
end
```

2. Compute the Cholesky factorization:

```
Q = chol(K, 'lower');
```

3. Generate realizations of the Gaussian process:

```
x = randn(n, 1); % Standard normal latent variables  
y = Q * x;      % Gaussian process realization
```

4. Analyze the columns or rows of Q to extract the impulse response:

```
h = Q(100:end, 100); % Extract a representative column
```

9 Conclusion

The Cholesky decomposition of a Gaussian process's covariance matrix provides a powerful causal interpretation. It reveals that:

- A Gaussian process can be viewed as the output of a causal linear time-invariant filter applied to white noise
- The filter's impulse response is encoded in the columns of the Cholesky factor
- For stationary processes, this impulse response converges to a fixed pattern
- The process can be generated sequentially, with each new observation depending only on the current and past latent variables

This interpretation establishes a bridge between Gaussian processes and linear systems theory, providing insights into both fields. It demonstrates how the rich covariance structure of a Gaussian process can be decomposed into simple, independent sources of randomness through a causal filtering operation.

This causal perspective is particularly valuable in applications requiring prediction, sequential data generation, or real-time processing of Gaussian process models.