# Lecture 11: GMRF vs. AR Models

Chapter 4.5–4.6

Charles A. Bouman

Purdue University

Friday, September 18, 2026

# Outline

1. **Comparing AR and GMRF models**
   The autocovariance equations, and the relationship between the two sets of model parameters.
2. **Equivalence in 1-D**
   The equivalence of the two models in one dimension, and the AR(1) example.
3. **What changes in 2-D**
   The growth of the prediction window, and the failure of factorization.
4. **GMRFs on general lattices**
   Neighborhood systems, pairwise cliques, and the density function.

Reading: Chapter 4.5–4.6, *Foundations of Computational Imaging*.

# Today's Big Ideas

1. In 1-D, an AR model and a GMRF of equal order are the same processes.
2. AR to GMRF is autocorrelation; GMRF to AR is factorization.
3. In 2-D, factorization fails, so GMRFs are more general than AR.
4. On any lattice, neighbors are pairs with nonzero off-diagonal B.

# Comparing AR and GMRF models in 1-D

Which model is more general, the AR model or the GMRF? A zero-mean stationary Gaussian process is completely specified by its autocovariance. So if the stationary AR and GMRF processes have the same autocovariance, then they have the same distribution, and the whole comparison reduces to the two autocovariance equations below.

| | 1-D AR Model | 1-D GMRF Model |
|---|---|---|
| Model Parameters | $\sigma_C^2$ , $h_n$ | $\sigma_{NC}^2$ , $g_n$ |
| Time Autocorrelation | $R_X(n) * (\delta_n - h_n) * (\delta_n - h_{-n}) = \sigma_C^2 \delta_n$ | $R_X(n) * (\delta_n - g_n) = \sigma_{NC}^2 \delta_n$ |
| Power Spectrum | $S_X(\omega) = \dfrac{\sigma_C^2}{\lvert 1 - H(\omega) \rvert^2}$ | $S_X(\omega) = \dfrac{\sigma_{NC}^2}{1 - G(\omega)}$ |

- The AR autocorrelation equation contains the prediction filter and its time reverse, while the GMRF equation contains its filter only once. The same difference appears in the power spectra, as $|1 - H|^2$ versus $1 - G$.

# The equations look the same in 1-D and 2-D

| | 2-D AR Model | 2-D GMRF Model |
|---|---|---|
| Model Parameters | $\sigma_C^2$ , $h_s$ | $\sigma_{NC}^2$ , $g_s$ |
| Space Autocorrelation | $R_X(s) * (\delta_s - h_s) * (\delta_s - h_{-s}) = \sigma_C^2 \delta_s$ | $R_X(s) * (\delta_s - g_s) = \sigma_{NC}^2 \delta_s$ |
| Power Spectrum | $S_X(\mu, \nu) = \dfrac{\sigma_C^2}{\lvert 1 - H(\mu, \nu) \rvert^2}$ | $S_X(\mu, \nu) = \dfrac{\sigma_{NC}^2}{1 - G(\mu, \nu)}$ |

The 2-D equations are formally identical to the 1-D equations, with $n$ replaced by $s = (s_1, s_2)$ and the DTFT replaced by the DSFT.

So why does the equivalence of the two models hold in 1-D but fail in 2-D? The difference is not in the convolution. It is in the **factorization**.

# Matching the two autocovariance equations

If we assume $R_X(n)$ is the same in both autocovariance equations, combining them gives

$$R_X(n) * (\delta_n - g_n)\,\sigma_C^2 = R_X(n) * (\delta_n - h_n) * (\delta_n - h_{-n})\sigma_{NC}^2 \ ,$$

which is satisfied when the model parameters obey

$$\sigma_C^2(\delta_n - g_n) = \sigma_{NC}^2(\delta_n - h_n) * (\delta_n - h_{-n}) \ . \tag{1}$$

GMRF side
$\delta_n - g_n$: symmetric

$\Longleftrightarrow$

AR side
$(\delta_n - h_n) * (\delta_n - h_{-n})$: causal $\times$ anticausal

AR $\rightarrow$ GMRF requires **autocorrelation**.
GMRF $\rightarrow$ AR requires **factorization**.

# Converting AR parameters to GMRF parameters

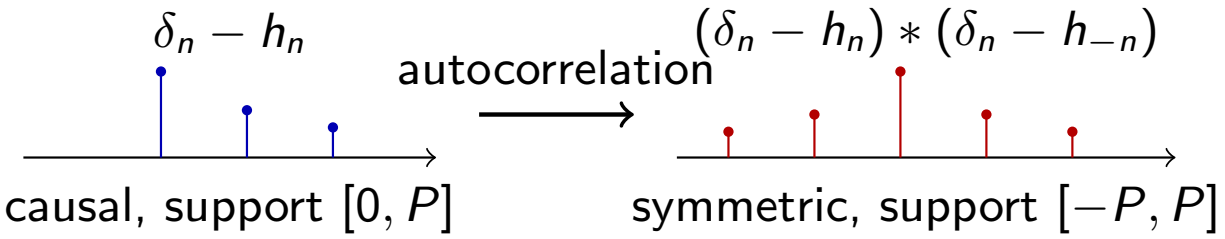
A causal AR predictor becomes a symmetric GMRF predictor by autocorrelation. Evaluating equation (1) at $n = 0$ gives the relationship between the prediction variances,

$$\sigma_{NC}^2 = \frac{\sigma_C^2}{1 + \sum_{n=1}^{P} h_n^2} ,$$

and using this relationship,

$$g_n = \delta_n - \frac{(\delta_n - h_n) * (\delta_n - h_{-n})}{1 + \sum_{n=1}^{P} h_n^2} .$$

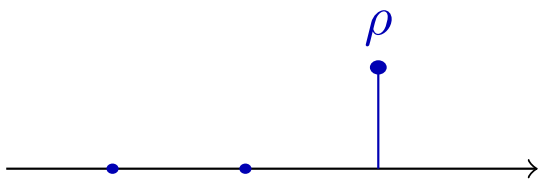
The **time autocorrelation** of a function $f_n$ is defined as $f_n * f_{-n}$, the convolution of the function with its time reverse. So the GMRF prediction coefficients are the normalized time autocorrelation of $\delta_n - h_n$:



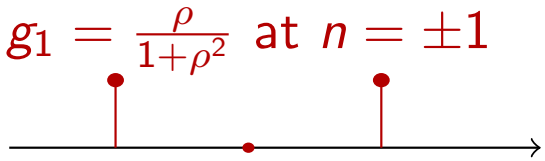
# Example 4.1: an AR(1) process as a GMRF

Consider a zero-mean stationary AR process of order $P = 1$. Its prediction filter is $h_n = \rho\,\delta_{n-1}$ with $|\rho| < 1$, and its prediction variance is $\sigma_C^2$. The parameters of the equivalent GMRF are then

$$\sigma_{NC}^2 = \frac{\sigma_C^2}{1+\rho^2}\ , \qquad g_n = \delta_n - \frac{(\delta_n - h_n) * (\delta_n - h_{-n})}{1+\rho^2} = \frac{\rho}{1+\rho^2}\,(\delta_{n-1} + \delta_{n+1})\ .$$




$h_n$: one-sided




$g_n$: symmetric

- Noncausal prediction variance is always smaller than causal.
- AR filter became symmetric, $g_n = g_{-n}$; model order did not grow.

# Example 4.1: an AR(1) process as a GMRF

The power spectrum can be computed from either parameterization. From the AR parameterization,

$$S_X(\omega) = \frac{\sigma_C^2}{|1 - H(\omega)|^2} = \frac{\sigma_C^2}{|1 - \rho e^{-j\omega}|^2} = \frac{\sigma_C^2}{(1+\rho^2)} \frac{1}{\left(1 - \frac{2\rho}{1+\rho^2}\cos(\omega)\right)},$$

and from the GMRF parameterization,

$$S_X(\omega) = \frac{\sigma_{NC}^2}{1 - G(\omega)} = \frac{\sigma_{NC}^2}{1 - 2g_1\cos(\omega)} = \frac{\sigma_C^2}{(1+\rho^2)} \frac{1}{\left(1 - \frac{2\rho}{1+\rho^2}\cos(\omega)\right)}.$$

The two expressions agree. This verifies that the AR model and the GMRF model give the same power spectral density.

# In 1-D the prediction window does not grow



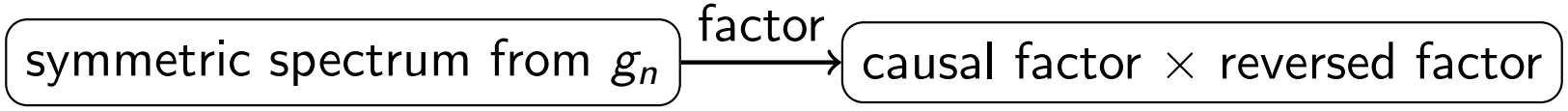

1-D AR order $P = 2$ 1-D GMRF order $P = 2$

- The pixels denoted by $\otimes$ are predicted using the values denoted by $\bullet$.
- Support $[0, P]$ becomes $[-P, P]$, so AR order $P$ gives GMRF order $P$.

# From GMRF parameters to AR parameters in 1-D

In order to find the parameters of an AR model from the parameters of a GMRF, it is necessary to find a causal predictor $h_n$ that satisfies equation (1),

$$\sigma_C^2(\delta_n - g_n) = \sigma_{NC}^2(\delta_n - h_n) * (\delta_n - h_{-n}) \ .$$



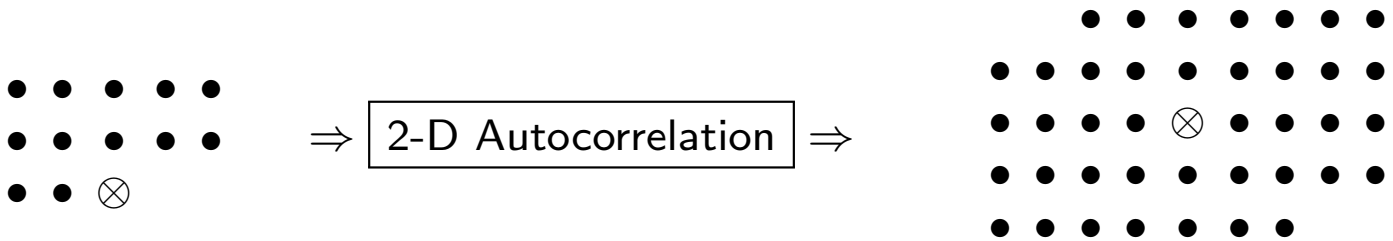

- In 1-D this is a classic Wiener filtering problem, already solved.
- Symmetric $g_n$: its Z-transform always factors causal $\times$ anticausal.

**Property: Equivalence of AR and GMRF models in 1-D**

A stationary discrete-time zero-mean Gaussian random process is a 1-D order-$P$ AR model **if and only if** it is a 1-D order-$P$ GMRF.

# In 2-D the prediction window grows

Does the same result hold for images? The question is what support the autocorrelation of a causal 2-D AR window has.




2-D AR order $P = 2$

Requires 2-D GMRF of order $2P = 4$

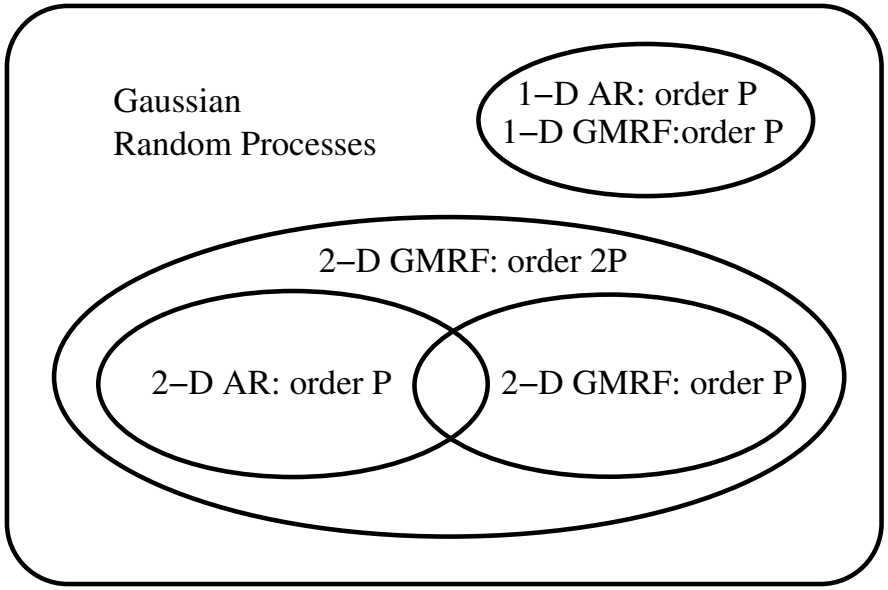
The 2-D convolution of the prediction filter with its reverse produces an oblong function which is $4P + 1$ wide and $2P + 1$ high.

**Property: In 2-D, an AR model of order $P$ is a GMRF of order $2P$**

In 2-D, a stationary discrete-time zero-mean order-$P$ AR model is also an order-$2P$ GMRF.

# The GMRF-to-AR factorization fails in 2-D

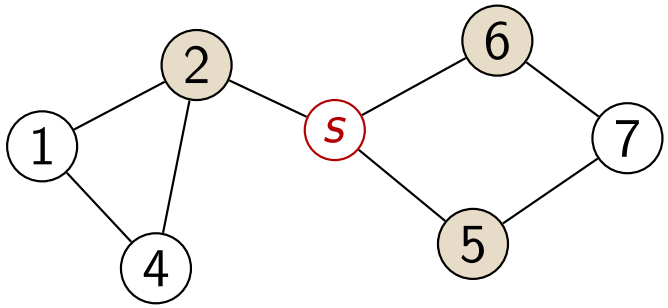
The factoring argument that worked in 1-D cannot be generalized to 2-D, because **polynomials in more than one dimension cannot, in general, be factored**. So in general a 2-D GMRF is not a 2-D AR process.




- In 1-D causality is a choice; in 2-D it excludes valid image models.
- This explains the name Gaussian Markov random **fields**: the concept grew out of a theory motivated by two or more dimensions, because that is where GMRFs are more general.

# A neighborhood system is an undirected graph

What counts as a neighborhood if the samples do not lie on a rectangular grid? Consider a random process $X_s$ indexed on a finite set of lattice points $s \in S$, with no assumption about the structure of $S$.




The shaded nodes are the neighbors $\partial s$ of node $s$.
The edges are undirected.

## Definition: Neighborhood System

A collection of sets $\partial s \subset S$, one for each $s \in S$, is a **neighborhood system** if it meets two conditions.
First, for all $s \in S$, $s \notin \partial s$.
Second, for all $s, r \in S$, if $s \in \partial r$, then $r \in \partial s$.

# Pairwise cliques, and the general definition of a GMRF

## Definition: Pairwise Clique

An unordered pair of pixels $\{s, r\}$ with $s, r \in S$ is said to be a **pairwise clique** if $s \in \partial r$. We denote the set of all pairwise cliques as

$$\mathcal{P} = \{ \{s, r\} : s \in \partial r \text{ such that } s, r \in S \} \ .$$

By convention, the pair $\{s, r\}$ is unordered, so each unique pair appears once in $\mathcal{P}$.

So each edge of the neighborhood graph is one pairwise clique.

## Definition: GMRF

A jointly Gaussian random process $X_s$ indexed on $s \in S$ is a **GMRF** with neighborhood system $\partial s$ if for all $s \in S$,

$$\mathbb{E}[X_s | X_r \text{ for } r \neq s] = \mathbb{E}[X_s | X_r \text{ for } r \in \partial s] \ .$$

So the noncausal predictor of $X_s$ depends only on its neighbors.

# Noncausal prediction on a general lattice

The equations of noncausal prediction carry over with the index set $S$ in place of $1, \cdots, N$:

$$\begin{aligned} \mathcal{E}_s &= X_s - \sum_{r \in \partial s} g_{s,r} X_r \ , \\ \sigma_s^2 &= \mathbb{E}\left[\mathcal{E}_s^2\right] \ , \end{aligned}$$

and in vector-matrix form,

$$\begin{aligned} \mathcal{E} &= (I - G)X \ , \\ \Gamma &= \text{diag}\left(\mathbb{E}\left[\mathcal{E}\mathcal{E}^t\right]\right) \ . \end{aligned}$$

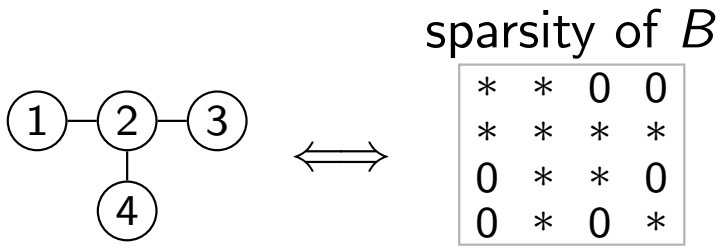
# The precision matrix encodes the neighborhood graph

## Property: Density of a zero-mean GMRF

Let $X$ be a zero-mean GMRF with noncausal prediction matrix $G$ and noncausal prediction variance matrix $\Gamma$. Then its density is

$$p(x) = \frac{1}{(2\pi)^{N/2}} |B|^{1/2} \exp\left\{ -\frac{1}{2} x^t B x \right\} ,$$

where $B = \Gamma^{-1}(I - G)$.




- Sparsity: $B_{s,r} = 0$ unless $\{s, r\}$ is a pairwise clique.

# Converting between the two descriptions of a GMRF

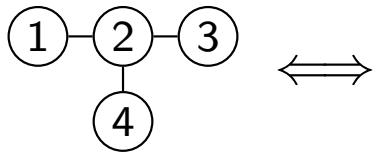
The prediction parameters $(\Gamma, G)$ and the inverse covariance matrix $B$ describe the same GMRF. Each one determines the other.

## $(\Gamma, G) \Rightarrow B$

$$B = \Gamma^{-1}(I - G)$$

## $B \Rightarrow (\Gamma, G)$

$$\Gamma^{-1} = \text{diag}(B)\ , \qquad G = I - \Gamma B$$




$\Longleftrightarrow$

sparsity of $B$

$$\begin{bmatrix} * & * & 0 & 0 \\ * & * & * & * \\ 0 & * & * & 0 \\ 0 & * & 0 & * \end{bmatrix}$$

- $g_{s,r}$ and $B_{s,r}$ differ only by the row scale $\sigma_s^2$, so sparsity matches.

# Summary

1. In 1-D, AR and GMRF are two parameterizations of the same stationary Gaussian models, because the two autocovariance equations agree when $\sigma_C^2(\delta_n - g_n) = \sigma_{NC}^2(\delta_n - h_n) * (\delta_n - h_{-n})$.
2. Going from AR to GMRF is autocorrelation, with $\sigma_{NC}^2 = \sigma_C^2/(1 + \sum_{n=1}^{P} h_n^2)$ and $g_n = \delta_n - (\delta_n - h_n) * (\delta_n - h_{-n})/(1 + \sum_{n=1}^{P} h_n^2)$. Going from GMRF to AR is factorization of the Z-transform of $\delta_n - g_n$.
3. That factorization fails in general in 2-D, where the autocorrelation of an order-$P$ window is $(4P+1)$ wide and $(2P+1)$ high, so an order-$P$ AR model is an order-$2P$ GMRF, but a 2-D GMRF is in general not a 2-D AR process.
4. A GMRF on any lattice is a graph model, with density determined by $B = \Gamma^{-1}(I - G)$, where $B_{s,r} = 0$ for $r \neq s$ unless $\{s, r\}$ is a pairwise clique.

Reading: Chapter 4.5–4.6. Next lecture: Chapter 5.1, MAP image restoration.