


# Conditioning in Gaussians

**Mark van der Wilk**

Department of Computing  
Imperial College London

@markvanderwilk  
m.vdwilk@imperial.ac.uk

November 16, 2021

## Last time

- ▶ Discussed the principle of Bayesian inference
- ▶ Skill: Compute posteriors
  - ▶ Example: Coin flipping (continuous latent, discrete obs)
  - ▶ Example: Noisy communication (discrete latent, continuous obs)
- ▶ Introduced: Noisy communication with Gaussian prior on signal.
- ▶ Skill: Gaussian conditioning.

To help with clarity, I made a notebook of figures to clarify last Friday's lecture (`110-bayes-voltage.ipynb`). Do carefully look at it.

# Gaussian Conditioning

Many important models are **Linear Gaussian Models**.

- ▶ Bayesian Linear Regression
- ▶ Factor Analysis / PCA / Linear Autoencoders
- ▶ Kalman filters (time series)

Special because:

- ▶ You can compute posteriors in closed form

(**closed form** means that you can express some solution in terms of typical functions log, sin, etc...)

- ▶ This is rare in Bayesian inference!

Key skill: Conditioning in Gaussian distributions.

## Example: Noisy analogue communication

Similar set-up to before.

$$p(s) = \mathcal{N}(s; 0, 1) \quad \text{Assume a prior on source voltage.} \quad (1)$$

$$p(v|s) = \mathcal{N}(v; s, \sigma^2) \quad \text{Assumptions about the channel.} \quad (2)$$

We are interested in our posterior belief on  $S$ :

$$p(s|v) = \frac{p(v|s)p(s)}{p(v)} \quad (3)$$

$$= \frac{p(v, s)}{\int p(v, s) ds} \quad \text{Alternative formulation (sum/prod rules)} \quad (4)$$

# Two formulations, two methods of solving

Both share the same goal:

1. Find an expression of the posterior density that you can directly implement (closed-form)
2. Express posterior as a known standard type of distribution

# Method 1: Crunching through densities

One method always works: Crunching through the densities.

We only care about terms that depend on  $s$ , since we know that a distribution has to normalise to 1!

$$p(s|v) \propto p(v|s)p(s) \tag{5}$$

$$= \mathcal{N}(v; s, \sigma^2) \mathcal{N}(s; 0, 1) \tag{6}$$

$$\propto \exp\left(-\frac{(v-s)^2}{2\sigma^2}\right) \exp\left(-\frac{s^2}{2}\right) \tag{7}$$

$$= \exp\left(-\frac{v^2}{2\sigma^2} + \frac{sv}{\sigma^2} - \frac{s^2}{2\sigma^2} - \frac{s^2}{2}\right) \tag{8}$$

$$\propto \exp\left(-\frac{1+\sigma^2}{2\sigma^2}s^2 + \frac{v}{\sigma^2}s\right) \tag{9}$$

## Method 1: Equating coefficients

$$\mathcal{N}(s; a, b) = c \cdot \exp\left(-\frac{1}{2b}x^2 + \frac{a}{b}x\right) \quad (10)$$

$$\implies b = \frac{\sigma^2}{1 + \sigma^2} \quad (11)$$

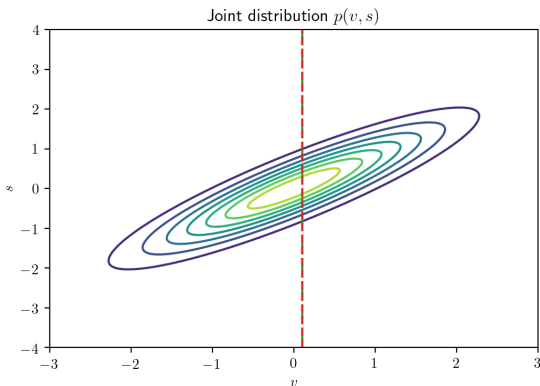
$$\implies a = \frac{1}{1 + \sigma^2}v \quad (12)$$

$$\implies p(s|v) = \mathcal{N}\left(s; \frac{1}{1 + \sigma^2}v, \frac{\sigma^2}{1 + \sigma^2}\right) \quad (13)$$

- ▶ Best guess (mode/mean) is more conservative than MaxLik (biased towards the mean of the prior)
- ▶ Variance tends to zero as  $\sigma^2 \rightarrow 0$

## Method 2: Joint and Conditioning Formula

Alter: we could find the joint first, and then evaluate along a line





## Method 2: Gaussian Conditioning Formula

For a joint Gaussian density

$$p(x, y) = \mathcal{N}\left(\begin{bmatrix} x \\ y \end{bmatrix}; \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} A & B \\ B & C \end{bmatrix}\right) \quad (14)$$

we have the conditional

$$p(y|x) = \mathcal{N}\left(y; \frac{B}{A}(x - a) + b, C - \frac{B^2}{A}\right) \quad (15)$$

# Linear Gaussian Model

Linear-Gaussian model:

- ▶ All RVs are conditionally Gaussian.
- ▶ Conditional means can depend **linearly** on other RVs.
- ▶ Can always rewrite as linear combination of independent Gaussian RVs.
- ▶ Linear transformations of Gaussians are still Gaussian (earlier lectures / exercise)
- ▶ This gives an easy way of finding the joint.

## Method 2: Example

Set-up:

$$p(s) = \mathcal{N}(s; 0, 1) \quad \text{Assume a prior on source voltage.} \quad (16)$$

$$p(v|s) = \mathcal{N}(v; s, \sigma^2) \quad \text{Assumptions about the channel.} \quad (17)$$

Write as linear transform, use expectation identities to find means and covariances:

$$\begin{bmatrix} s \\ v \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} s \\ \epsilon \end{bmatrix} \quad (18)$$

$$\mathbb{E}_{s,\epsilon}[v] = \mathbb{E}_{s,\epsilon}[s + \epsilon] = \mathbb{E}_s[s] + \mathbb{E}_\epsilon[\epsilon] = 0 \quad (19)$$

$$\mathbb{V}_{s,\epsilon}[v] = \mathbb{V}_{s,\epsilon}[s + \epsilon] = \mathbb{V}_s[s] + \mathbb{V}_\epsilon[\epsilon] = 1 + \sigma^2 \quad \text{Indep.} \quad (20)$$

$$\mathbb{C}_{s,\epsilon}[s, v] = \mathbb{E}_{s,\epsilon}[(s + \epsilon)s] = \mathbb{E}_s[s]\mathbb{E}_\epsilon[\epsilon] + \mathbb{E}_\epsilon[s^2] \quad \text{Indep.} \quad (21)$$

$$= 0 + \mathbb{V}_s[s] = 1 \quad \mathbb{E}_\epsilon[\epsilon] = 0 \quad (22)$$

## Method 2: Example continued

Given the means, variances and covariances we have computed, we can fill in the full joint:

$$p(s, v) = \mathcal{N}\left(\begin{bmatrix} v \\ s \end{bmatrix}; \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 + \sigma^2 & 1 \\ 1 & 1 \end{bmatrix}\right) \quad (23)$$

We can now apply the Gaussian conditioning formula (make sure  $s, v$  are the right way round):

$$p(s|v) = \mathcal{N}\left(s; \frac{\sigma^2}{1 + \sigma^2}v, 1 - \frac{1}{1 + \sigma^2}\right) \quad (24)$$

$$= \mathcal{N}\left(s; \frac{\sigma^2}{1 + \sigma^2}v, \frac{\sigma^2}{1 + \sigma^2}\right) \quad (25)$$

# Conclusion

Gaussian conditioning:

- ▶ Method 1: Equating coefficients
- ▶ Method 2: Finding joint, using conditioning formula