# DATA ASSIMILATION REPORT

JULY 18, 2022

MALAK HAMED MOKHTAR
Student Number: 612208004

Proof of the lemma: $\pi_{X^1}(x^1) = \mathbb{E}[\pi_{X^1}(x^1|X^2)]$ , where $X^1$ and $X^2$ are two random variables with joint PDF $\pi_{X^1X^2}$ .

$\because$ If $\wedge^j = \mathbb{R}^{Nj}$ and if $X^j$ is absolutely continuous, then $Z = (X^1, X^2)$ is absolutely continuous, with pdf denoted by $\pi_{X^1X^2}$ and one has:

The marginal PDFs:

$$\pi_{X^1}(x^1) = \int_{\mathbb{R}^{N_2}} \pi_{X^1X^2}(x^1, x^2)dx^2 \ . \rightarrow \quad 1$$

And

$$\pi_{X^2}(x^2) = \int_{\mathbb{R}^{N_1}} \pi_{X^1X^2}(x^1, x^2)dx^1 \ . \rightarrow \quad 2$$

Additionally, one has the disintegration formula:

$$\pi_{X^n,\dots,X^1}(x^n, \dots, x^1) = \pi_{X^n}(x^n|x^{n-1}, \dots, x^1)\pi_{X^{n-1}}(x^{n-1}|x^{n-2}, \dots, x^1) \dots \pi_{X^2}(x^2|x^1)\pi_{X^1}(x^1). \rightarrow \ 3$$

For two random variables we get:

$$\pi_{X^2,X^1}(x^2, x^1) = \pi_{X^2}(x^2|x^1)\pi_{X^1}(x^1) \ . \rightarrow \quad 4$$

$\because$ $\pi_{X^2,X^1}(x^2, x^1) = \pi_{X^1,X^2}(x^1, x^2)$

$\therefore$ (4) can be re-written as: $\pi_{X^1,X^2}(x^1, x^2) = \pi_{X^1}(x^1|x^2)\pi_{X^2}(x^2) \ . \rightarrow \quad 5$

Substituting (5) in (1) we get:

$\pi_{X^1}(x^1) = \int_{\mathbb{R}^{N_2}} \pi_{X^1}(x^1|x^2)\pi_{X^2}(x^2) \ dx^2$

$= \mathbb{E}[\pi_{X^1}(x^1|X^2)]$, where the expectation is taken with respect to the random variable $X^2$.

Proof of Conditional distributions of the multivariate normal distribution:

Assuming $X$ follows a multivariate normal distribution of $X \sim \mathcal{N}(\bar{x}, \sigma)$ , with mean $\left(\overline{x^1}, \overline{x^2}\right)$, covariance matrix: $\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$. Using the convention $\sigma$ instead of $\sigma^2$ for simplification. This would mean that the conditional distribution of any subset vector $X^1$, given the complement vector $X^2$, is also a multivariate normal distribution:

$$X^1|X^2 \sim \mathcal{N}\left(\overline{x^{1|2}}, \sigma_{1|2}\right).$$

With the mean:

$$\bar{x}^{1|2} = x^1 + \sigma_{12}\sigma_{22}^{-1}\left(x^2 - \overline{x^2}\right),$$

and the covariance:

$$\sigma_{1|2} = \sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21}.$$

Proof, without loss of generality, assuming $x = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}$ where $x^1$ $is\ a\ n^1 \times 1$ vector, $x^2$ $is\ a\ n^2 \times 1$ vector and $x$ $is\ a\ n^1 + n^2 = n \times 1$ vector.

By construction, the joint distribution of $X^1$ $and$ $X^2$ is:

$$X^1, X^2 \sim \mathcal{N}(\bar{x}, \sigma).$$

As for the marginal distribution of $X^2$ one can obtain from the assumptions that:

$$X^2 \sim \mathcal{N}\left(\bar{x}^2, \sigma_{22}\right).$$

From the conditional probability distribution function we have:

$$\pi(x^1\,|x^2) = \frac{\pi(x^1, x^2)}{\pi(x^2\,)} = \frac{\mathcal{N}(x; \bar{x}, \sigma)}{\mathcal{N}(x^2; \bar{x}^2, \sigma_{22})}\,. \rightarrow \; 1$$

Applying the probability density function of the multivariate normal distribution (i.e. from (1)):

$$\pi(x^1\,|x^2) = \frac{\frac{1}{\sqrt{(2\pi)^n}|\sigma|^{\frac{1}{2}}} e^{\frac{-1}{2}(x-\bar{x})^T\sigma^{-1}(x-\bar{x})}}{\frac{1}{\sqrt{(2\pi)^{n^2}}|\sigma_{22}|^{\frac{1}{2}}} e^{\frac{-1}{2}\left(x^2-\overline{x^2}\right)^T\sigma_{22}{}^{-1}\left(x^2-\overline{x^2}\right)}}$$

$$= \frac{1}{\sqrt{(2\pi)^{n-n^2}}} \cdot \sqrt{\frac{|\sigma_{22}|}{|\sigma|}} \,.\, e^{\frac{-1}{2}(x-\bar{x})^T\sigma^{-1}(x-\bar{x})+\frac{1}{2}\left(x^2-\overline{x^2}\right)^T\sigma_{22}{}^{-1}\left(x^2-\overline{x^2}\right)} \;\; \rightarrow \;\; 2$$

where,

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}, \quad \rightarrow 3$$

And writing the inverse as:

$$\sigma^{-1} = \begin{bmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{bmatrix}. \quad \rightarrow 4$$

Substituting (3) & (4) in (2):

$$\pi(x^1 | x^2) = \frac{1}{\sqrt{(2\pi)^{n-n^2}}} \cdot \sqrt{\frac{|\sigma_{22}|}{|\sigma|}} \cdot e^{\frac{-1}{2}\left(\begin{bmatrix} x^1 \\ x^2 \end{bmatrix} - \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}\right)^T \begin{bmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{bmatrix} \left(\begin{bmatrix} x^1 \\ x^2 \end{bmatrix} - \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}\right) + \frac{1}{2}(x^2 - \overline{x^2})^T \sigma_{22}{}^{-1}(x^2 - \overline{x^2})}.$$

Multiplying out within the exponent one can obtain:

$$\pi(x^1 | x^2) = \frac{1}{\sqrt{(2\pi)^{n-n^2}}} \cdot \sqrt{\frac{|\sigma_{22}|}{|\sigma|}} \cdot$$
$$e^{\frac{-1}{2}\left((x^1 - \overline{x^1})^T \sigma^{11}(x^1 - \overline{x^1}) + 2(x^1 - \overline{x^1})^T \sigma^{12}(x^2 - \overline{x^2}) + (x^2 - \overline{x^2})^T \sigma^{22}(x^2 - \overline{x^2})\right) + \frac{1}{2}(x^2 - \overline{x^2})^T \sigma_{22}^{-1}(x^2 - \overline{x^2})} . \rightarrow 5$$

Generally,

The inverse of a block matrix is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$$

Therefore making use of the above property and substituting in (5) one can obtain:

$$\pi(x^1 | x^2) = \frac{1}{\sqrt{(2\pi)^{n-n^2}}} \cdot \sqrt{\frac{|\sigma_{22}|}{|\sigma|}} \cdot$$
$$e^{\frac{-1}{2}\left(\begin{matrix}(x^1 - \overline{x^1})^T (\sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21})^{-1}(x^1 - \overline{x^1}) - 2(x^1 - \overline{x^1})^T (\sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21})^{-1} \sigma_{12}\sigma_{22}^{-1}(x^2 - \overline{x^2}) + \\ (x^2 - \overline{x^2})^T [\sigma_{22}^{-1} + \sigma_{21}\sigma_{22}^{-1}(\sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21})^{-1} \sigma_{12}\sigma_{22}^{-1}](x^2 - \overline{x^2})\end{matrix}\right) + \frac{1}{2}(x^2 - \overline{x^2})^T \sigma_{22}^{-1}(x^2 - \overline{x^2})} . \rightarrow 6$$

Eliminating some terms in (6), rearranging them and using the fact that $\sigma$ is a covariance matrix (meaning that $\sigma_{21} = \sigma_{12}^T$), will get us:

$$\pi(x^1 | x^2) =$$
$$\frac{1}{\sqrt{(2\pi)^{n-n^2}}} \cdot \sqrt{\frac{|\sigma_{22}|}{|\sigma|}} \cdot e^{\frac{-1}{2}\left[x^1 - \left(\overline{x^1} + \sigma_{12}\sigma_{22}^{-1}(x^2 - \overline{x^2})\right)\right]^T (\sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21})^{-1} \left[x^1 - \left(\overline{x^1} + \sigma_{12}\sigma_{22}^{-1}(x^2 - \overline{x^2})\right)\right]} . \rightarrow 7$$

Generally,

The determinant of a block matrix is

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \cdot |A - BD^{-1}C|$$

Therefore making use of the above property as well as that $n^1 = n - n^2$ and substituting in (7), we can finally conclude that:

$$\pi(x^1 \,|x^2) = \frac{1}{\sqrt{(2\pi)^{n^1}|\sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21}|}} \, . \, e^{\frac{-1}{2}\left[x^1-\left(\overline{x^1}+\sigma_{12}\sigma_{22}^{-1}(x^2-\overline{x^2})\right)\right]^T(\sigma_{11}-\sigma_{12}\sigma_{22}^{-1}\sigma_{21})^{-1}\left[x^1-\left(\overline{x^1}+\sigma_{12}\sigma_{22}^{-1}(x^2-\overline{x^2})\right)\right]} \, .$$

Which is the probability density function of a multivariate normal distribution:

$$\pi(x^1 \,|x^2) = \mathcal{N}\left(x^1; \bar{x}^{1|2}, \sigma_{1|2}\right).$$

Example 2.23

(In the previous example the convention $\sigma$ instead of $\sigma^2$ was used, however, for this example the convention $\sigma^2$ will be used instead to additionally explore dealing with $\sigma^2$.

Setting $N_x = 1$ and considering a pair of univariate Gaussian PDFs:

$$\pi_{X^i}(x^i) = \mathcal{N}(x^i; \overline{x^i}, \sigma_{ii}^2)\text{ , where } i = 1{,}2.$$

One can seek a coupling in $z = (x^1, x^2)$ of the form:

$$\pi_Z(z) = \mathcal{N}(z; \bar{z}, P).$$

Therefore, using the given marginals one can obtain that:

$$\bar{z} = (\overline{x^1}, \overline{x^2})^T,$$

and that the covariance matrix P must take the form:

$$P = \begin{bmatrix} \sigma_{11}^2 & \rho\sigma_{11}\sigma_{22} \\ \rho\sigma_{11}\sigma_{22} & \sigma_{22}^2 \end{bmatrix}.$$

Since $P$ has to be positive definite then the following correlation:

$$\rho = \frac{\sigma_{12}^2}{|\sigma_{11}||\sigma_{22}|} = \frac{\sigma_{21}^2}{|\sigma_{11}||\sigma_{22}|},$$

between $X^1$ and $X^2$ must satisfy: $\rho^2 \leq 1$

The PDF $\pi_{X_i}(x)$ of a Gaussian random variable $X_i \sim \mathrm{N}(\overline{x}_i, \sigma_{ii}^2)$ is denoted by $\mathrm{n}(x; \overline{x}_i, \sigma_{ii}^2)$. For given $x_2$, we define $X_1|x_2$ as the random variable with conditional probability distribution $\pi_{X_1}(x_1|x_2)$. From the previous calculations we then obtain $X_1|x_2 \sim \mathrm{N}(\overline{x}_c, \sigma_c^2)$ and

$$\pi_{X_1X_2}(x_1, x_2) = \mathrm{n}(x_1; \overline{x}_c, \sigma_c^2)\, \mathrm{n}(x_2; \overline{x}_2, \sigma_{22}^2). \tag{2.8}$$

Further exploring the conditional PDF:

$$\pi_{X^2}(x^2|x^1).$$

And making use of the following, one can obtain:

$$\pi_{X^2}(x^2|x^1) = \mathcal{N}(x^2; \overline{x^c}, \sigma_c^2),$$

where,

$$\overline{x^c} = \overline{x^2} - \frac{\sigma_{21}^2}{\sigma_{11}^2}\left(\overline{x^1} - x^1\right) = \overline{x^2} - \rho\frac{\sigma_{22}}{\sigma_{11}}\left(\overline{x^1} - x^1\right) \text{ and } \sigma_c^2 = \sigma_{22}^2 - \sigma_{12}^4\sigma_{11}^{-2} = (1-\rho^2)\sigma_{22}^2$$

As $\rho \to 1$, $\sigma_c \to 0$ implying that the conditional probability density becomes a Dirac Delta distribution centered about:

$$\overline{x^c} = \overline{x^2} - \frac{\sigma_{22}}{\sigma_{11}}\left(\overline{x^1} - x^1\right).$$

Therefore, optimizing the correlation between $X^1$ and $X^2$ has led us to a deterministic coupling:

$$X^2 = \overline{x^2} + \frac{\sigma_{22}}{\sigma_{11}}\left(X^1 - \overline{x^1}\right). \rightarrow 1$$

As outlined in the reference book “Probabilistic Forecasting and Bayesian Data Assimilation” by “Sebastian Reich and Colin Cotter”. The above coupling is “commonly used in order to transform random numbers from a $\mathrm{N}(0, 1)$ random number generator into Gaussian random numbers with mean $\bar{x}$ and variance $\sigma^2$. This particular case leads to $\overline{x^1} = 0$ , $\overline{x^2} = \bar{x}$ , $\sigma_{11} = 1$ $and$ $\sigma_{22} = \sigma$ in (1). We will find later that this coupling is optimal in the sense that it maximizes the correlation between $X^1$ and $X^2$. Joint distributions in $z$ are shown in the below Figure for a small correlation $\rho = 0.1$ and a nearly optimal correlation $\rho = 0.95$ in the second figure”.