CS329 Machine Learning

Homework #2

Fan Site
fanst2021@mail.sustech.edu.cn

Question 1

(a) **[True or False]** If two sets of variables are jointly Gaussian, then the conditional distribution of one set conditioned on the other is again Gaussian. Similarly, the marginal distribution of either set is also Gaussian.

(b) Consider a partitioning of the components of into three groups x_a, x_b, and x_c, with a corresponding partitioning of the mean vector μ and of the covariance matrix $\mathbf{\Sigma}$ in the form

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \\ \mu_c \end{pmatrix} \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{aa} & \mathbf{\Sigma}_{ab} & \mathbf{\Sigma}_{ac} \\ \mathbf{\Sigma}_{ba} & \mathbf{\Sigma}_{bb} & \mathbf{\Sigma}_{bc} \\ \mathbf{\Sigma}_{ca} & \mathbf{\Sigma}_{cb} & \mathbf{\Sigma}_{cc} \end{pmatrix}$$

Find an expression for the conditional distribution $p(x_a|x_b)$ in which x_c has been marginalized out.

Solution (a)

True.

1. Conditional

Given $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{\Sigma})$, partitioned as $\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \end{pmatrix}$, $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix}$, $\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{aa} & \mathbf{\Sigma}_{ab} \\ \mathbf{\Sigma}_{ba} & \mathbf{\Sigma}_{bb} \end{pmatrix}$.

Assume that $\mathbf{x}_a = \mathbf{A}\mathbf{x}_b + \mathbf{w}$, $\mathbf{\Sigma}_{a|b} = \mathbf{\Sigma}_{\mathbf{w}}$.

We have $\mathbf{x}_a - \mu_a = \mathbf{A}(\mathbf{x}_b - \mu_b) + \mathbf{w}$, therefore $\mu_{a|b} - \mu_a = \mathbf{A}(\mathbf{x}_b - \mu_b)$.

From $\mathbf{\Sigma}_{ab} = \mathbf{A}\mathbf{\Sigma}_{bb}$ we obtain that $A = \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}$.

So we have

$$\mathbf{\Sigma}_{aa} = \mathbf{A}\mathbf{\Sigma}_{bb}\mathbf{A}^{\mathrm{T}} + \mathbf{\Sigma}_{\mathbf{w}} = \mathbf{A}\mathbf{\Sigma}_{bb}\mathbf{A}^{\mathrm{T}} + \mathbf{\Sigma}_{a|b} = \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba} + \mathbf{\Sigma}_{a|b}$$

Then solve for $\mu_{a|b}$ and $\mathbf{\Sigma}(a|b)$

$$\mu_{a|b} = \mu_a + \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}(\mathbf{x}_b - \mu_b)$$

$$\mathbf{\Sigma}_{a|b} = \mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba}$$

So the conditional distribution of one set conditioned on the other is again Gaussian.

2. Marginal

$$p(\mathbf{x}_a) = \int p(\mathbf{x}_a, \mathbf{x}_b)\mathrm{d}\mathbf{x}_b = \mathcal{N}(\mathbf{x}_a|\boldsymbol{\mu}_a, \mathbf{\Sigma}_{aa})$$

So the marginal distribution of either set is also Gaussian.

Solution (b)

First, partition $\mathbf{x}$ as $\begin{pmatrix} \mathbf{x}_{a,b} \\ \mathbf{x}_c \end{pmatrix}$ and $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{a,b} \\ \boldsymbol{\mu}_c \end{pmatrix}$, $\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{(a,b)(a,b)} & \mathbf{\Sigma}_{(a,b)c} \\ \mathbf{\Sigma}_{c(a,b)} & \mathbf{\Sigma}_{cc} \end{pmatrix}$.

From the result of *Question (a): Marginal*, we have

$$p(\mathbf{x}_{a,b}) = \mathcal{N}\left(\mathbf{x}_{a,b}|\boldsymbol{\mu}_{a,b}, \mathbf{\Sigma}_{(a,b)(a,b)}\right)$$

Second, we partition again: $\boldsymbol{\mu}_{a,b} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix}$, $\mathbf{\Sigma}_{(a,b)(a,b)} = \begin{pmatrix} \mathbf{\Sigma}_{aa} & \mathbf{\Sigma}_{ab} \\ \mathbf{\Sigma}_{ba} & \mathbf{\Sigma}_{bb} \end{pmatrix}$.

From the resulf of *Question (a): Conditional*, we have

$$p(\mathbf{x}_a|\mathbf{x}_b) = \mathcal{N}\left(\mathbf{x}|\boldsymbol{\mu}_{a|b}, \boldsymbol{\Sigma}_{a|b}\right)$$

where

$$\mu_{a|b} = \mu_a + \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}(\mathbf{x}_b - \mu_b)$$

$$\boldsymbol{\Sigma}_{a|b} = \boldsymbol{\Sigma}_{aa} - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}\boldsymbol{\Sigma}_{ba}$$

Question 2

Consider a joint distribution over the variable $\mathbf{z} = \binom{\mathrm{x}}{\mathrm{y}}$ whose mean and covariance are given by

$$\mathbb{E}[\mathbf{z}] = \begin{pmatrix} \mu \\ \mathbf{A}\mu + \mathbf{b} \end{pmatrix} \quad \operatorname{cov}[\mathbf{z}] = \begin{pmatrix} \boldsymbol{\Lambda}^{-1} & \boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \\ \mathbf{A}\boldsymbol{\Lambda}^{-1} & \mathbf{L}^{-1} + \mathbf{A}\boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} \end{pmatrix}$$

(a) Show that the marginal distribution $p(\mathbf{x})$ is given by $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mu, \boldsymbol{\Lambda}^{-1})$.

(b) Show that the conditional distribution $p(\mathbf{y}|\mathbf{x})$ is given by $p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x}{+}\mathbf{b},\ \mathbf{L^{-1}})$.

Solution (a)

From the result of *Question (a): Marginal*, we have

$$p(\mathbf{x}) = \int p(\mathbf{x}, \mathbf{y})\mathrm{d}\mathbf{y} = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}_{aa}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1})$$

Solution (b)

From the resulf of *Question (a): Conditional*, we have

$$p(\mathbf{y}|\mathbf{x}) = \mathcal{N}\left(\mathbf{y}|\boldsymbol{\mu}_{y|x}, \boldsymbol{\Sigma}_{y|x}\right)$$

where

$$\boldsymbol{\mu}_{y|x} = \mathbf{A}\mu + \mathbf{b} + \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\mathbf{x} - \mu) = \mathbf{A}\mu + \mathbf{b} + \mathbf{A}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}(\mathbf{x} - \mu) = \mathbf{A}\mathbf{x} + \mathbf{b}$$

$$\boldsymbol{\Sigma}_{y|x} = \boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy} = \mathbf{L}^{-1} + \mathbf{A}\boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} - \mathbf{A}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}} = \mathbf{L}^{-1}$$

i.e.,

$$p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x}{+}\mathbf{b},\ \mathbf{L^{-1}})$$

Question 3

Show that the covariance matrix $\boldsymbol{\Sigma}$ that maximizes the log likelihood function is given by the sample covariance

$$\ln p(\mathbf{X}|\mu, \boldsymbol{\Sigma}) = -\frac{ND}{2}\ln(2\pi) - \frac{N}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{n=1}^{N}(\mathbf{x}_n - \mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \mu).$$

Is the final result symmetric and positive definite (provided the sample covariance is nonsingular)?

Solution

Derive the log likelihood function with respect to $\boldsymbol{\Sigma}$ and let it be 0

$$\frac{\partial \ln p(\mathbf{X}|\mu, \boldsymbol{\Sigma})}{\partial \boldsymbol{\Sigma}} = -\frac{\partial}{\partial \boldsymbol{\Sigma}}\left(\frac{N}{2}\ln|\boldsymbol{\Sigma}| + \frac{1}{2}\sum_{n=1}^{N}(\mathbf{x}_n - \mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \mu)\right) = 0$$

For the first term

$$-\frac{N}{2}\frac{\partial}{\partial \mathbf{\Sigma}}\ln|\mathbf{\Sigma}| = -\frac{N}{2}\left(\mathbf{\Sigma}^{-1}\right)^{\mathrm{T}} = -\frac{N}{2}\mathbf{\Sigma}^{-1}$$

For the second term

$$\begin{aligned}-\frac{1}{2}\sum_{n=1}^{N}\frac{\partial}{\partial \mathbf{\Sigma}}(\mathbf{x}_n-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}_n-\mu) &= \frac{1}{2}\sum_{n=1}^{N}\mathbf{\Sigma}^{-\mathrm{T}}(\mathbf{x}_n-\mu)(\mathbf{x}_n-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-\mathrm{T}} \\ &= \frac{1}{2}\mathbf{\Sigma}^{-1}\sum_{n=1}^{N}\left\{(\mathbf{x}_n-\mu)(\mathbf{x}_n-\mu)^{\mathrm{T}}\right\}\mathbf{\Sigma}^{-1}\end{aligned}$$

So we obtain

$$-\frac{N}{2}\mathbf{\Sigma}^{-1} + \frac{1}{2}\mathbf{\Sigma}^{-1}\sum_{n=1}^{N}\left\{(\mathbf{x}_n-\mu)(\mathbf{x}_n-\mu)^{\mathrm{T}}\right\}\mathbf{\Sigma}^{-1} = 0$$

$$\mathbf{\Sigma}_{\mathrm{ML}} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{x}_n-\mu)(\mathbf{x}_n-\mu)^{\mathrm{T}}$$

$$\mathbf{\Sigma}_{\mathrm{ML}}^{\mathrm{T}} = \mathbf{\Sigma}_{\mathrm{ML}},\ \text{For any } \mathbf{x}, \mathbf{x}^{T}\Sigma_{\mathrm{ML}}\mathbf{x} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}^{T}(\mathbf{x}_n-\mu)(\mathbf{x}_n-\mu)^{\mathrm{T}}\mathbf{x} = \frac{1}{N}\sum_{n=1}^{N}\left(\mathbf{x}^{T}(\mathbf{x}_n-\mu)\right)^2 > 0$$

therefore $\mathbf{\Sigma}_{\mathrm{ML}}$ is symmetric and positive-definite.

Question 4

(a) Derive an expression for the sequential estimation of the variance of a univariate Gaussian distribution, by starting with the maximum likelihood expression

$$\boldsymbol{\sigma}_{\mathrm{ML}}^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^2.$$

Verify that substituting the expression for a Gaussian distribution into the Robbins-Monro sequential estimation formula gives a result of the same form, and hence obtain an expression for the corresponding coefficients a_N.

(b) Derive an expression for the sequential estimation of the covariance of a multivariate Gaussian distribution, by starting with the maximum likelihood expression

$$\mathbf{\Sigma}_{\mathrm{ML}} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{x}_n-\mu_{\mathrm{ML}})(\mathbf{x}_n-\mu_{\mathrm{ML}})^{\mathrm{T}}.$$

Verify that substituting the expression for a Gaussian distribution into the Robbins-Monro sequential estimation formula gives a result of the same form, and hence obtain an expression for the corresponding coefficients a_N.

Solution (a)

$$
\begin{aligned}
(\sigma^2_{\text{ML}})^{(N)} &= \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu_{\text{ML}})^2 \\
&= \frac{1}{N}\sum_{n=1}^{N-1}(x_n-\mu_{\text{ML}})^2+\frac{1}{N}(x_N-\mu_{\text{ML}})^2 \\
&= \left(1-\frac{1}{N}\right)\frac{1}{N-1}\sum_{n=1}^{N-1}(x_n-\mu_{\text{ML}})^2+\frac{1}{N}(x_N-\mu_{\text{ML}})^2 \\
&= (\sigma^2_{\text{ML}})^{(N-1)}+\frac{1}{N}\left[(x_N-\mu_{\text{ML}})^2-(\sigma^2_{\text{ML}})^{(N-1)}\right]
\end{aligned}
$$

The maximum likelihood solution σ^2_{ML} satisfies

$$
\frac{\partial}{\partial\sigma^2}\left\{\frac{1}{N}\sum_{n=1}^{N}\ln p(x_n|\mu,\sigma^2)\right\}_{\sigma_{\text{ML}}}=0
$$

Exchanging the derivative and the sum, and taking the limit $N\to\infty$ we have

$$
\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\ln p(x_n|\mu,\sigma^2)=\mathbb{E}_x\left[\frac{\partial}{\partial\sigma^2}\ln p(x_n|\mu,\sigma^2)\right]
$$

Then we obtain the sequential formula to estimate σ_{ML}

$$
\begin{aligned}
(\sigma^2_{\text{ML}})^{(N)} &= (\sigma^2_{\text{ML}})^{(N-1)}+a_{N-1}\frac{\partial}{\partial(\sigma^2_{\text{ML}})^{(N-1)}}\ln p\left(x_n|\mu^N_{\text{ML}},(\sigma^2_{\text{ML}})^{(N-1)}\right) \\
&= (\sigma^2_{\text{ML}})^{(N-1)}+a_{N-1}\left[\frac{(x_N-\mu^N_{\text{ML}})^2}{2(\sigma^4_{\text{ML}})^{(N-1)}}-\frac{1}{2(\sigma^2_{\text{ML}})^{(N-1)}}\right]
\end{aligned}
$$

To make this equation equal to the first one, we have

$$
\begin{aligned}
a_{N-1} &= \frac{2}{N}(\sigma^4_{\text{ML}})^{(N-1)} \\
a_N &= \frac{2}{N+1}(\sigma^4_{\text{ML}})^{(N)}
\end{aligned}
$$

Solution (b)

$$
\begin{aligned}
\mathbf{\Sigma}^{(N)}_{\text{ML}} &= \frac{1}{N}\sum_{n=1}^{N}(\mathbf{x}_n-\boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_n-\boldsymbol{\mu}_{\text{ML}})^{\text{T}} \\
&= \frac{1}{N}\sum_{n=1}^{N-1}(\mathbf{x}_n-\boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_n-\boldsymbol{\mu}_{\text{ML}})^{\text{T}}+\frac{1}{N}(\mathbf{x}_N-\boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_N-\boldsymbol{\mu}_{\text{ML}})^{\text{T}} \\
&= \left(1-\frac{1}{N}\right)\frac{1}{N-1}\sum_{n=1}^{N-1}(\mathbf{x}_n-\boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_n-\boldsymbol{\mu}_{\text{ML}})^{\text{T}}+\frac{1}{N}(\mathbf{x}_N-\boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_N-\boldsymbol{\mu}_{\text{ML}})^{\text{T}} \\
&= \mathbf{\Sigma}^{(N-1)}_{\text{ML}}+\frac{1}{N}\left[(\mathbf{x}_N-\boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_N-\boldsymbol{\mu}_{\text{ML}})^{\text{T}}-\mathbf{\Sigma}^{(N-1)}_{\text{ML}}\right]
\end{aligned}
$$

We obtain the sequential formula to estimate $\mathbf{\Sigma}_{\text{ML}}$

$$
\begin{aligned}
\boldsymbol{\Sigma}_{\text{ML}}^{(N)} &= \boldsymbol{\Sigma}_{\text{ML}}^{(N-1)} + a_{N-1} \frac{\partial}{\partial \boldsymbol{\Sigma}_{\text{ML}}^{(N-1)}} \ln p\left(\mathbf{x}_N | \boldsymbol{\mu}_{\text{ML}}, \boldsymbol{\Sigma}_{\text{ML}}^{(N-1)}\right) \\
&= \boldsymbol{\Sigma}_{\text{ML}}^{(N-1)} + \frac{a_{N-1}}{2}\left[-(\boldsymbol{\Sigma}_{\text{ML}}^{-1})^{(N-1)} + (\boldsymbol{\Sigma}_{\text{ML}}^{-1})^{(N-1)}(\mathbf{x}_N - \boldsymbol{\mu}_{\text{ML}})(\mathbf{x}_N - \boldsymbol{\mu}_{\text{ML}})^{\mathrm{T}}(\boldsymbol{\Sigma}_{\text{ML}}^{-1})^{(N-1)}\right]
\end{aligned}
$$

To make this equation equal to the one above, we have

$$
\mathbf{a}_{N-1} = \frac{2}{N}(\boldsymbol{\Sigma}_{\text{ML}}^2)^{(N-1)}
$$

$$
\mathbf{a}_N = \frac{2}{N+1}(\boldsymbol{\Sigma}_{\text{ML}}^2)^{(N)}
$$

Question 5

Consider a D-dimensional Gaussian random variable $\mathbf{x}$ with distribution $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ in which the covariance $\boldsymbol{\Sigma}$ is known and for which we wish to infer the mean $\boldsymbol{\mu}$ from a set of observations $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N\}$. Given a prior distribution $p(\boldsymbol{\mu}) = \mathcal{N}(\boldsymbol{\mu}|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, find the corresponding posterior distribution $p(\boldsymbol{\mu}|\mathbf{X})$.

Lemma (From PRML 2.71)

The exponent in a general Gaussian distribution $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ can be written as

$$
-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = -\frac{1}{2}\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \text{const.}
$$

where const denotes terms which are independent of $\mathbf{x}$.

Solution

Posterior $\propto$ Prior $\times$ Likelihood, so

$$
\begin{aligned}
p(\boldsymbol{\mu}|\mathbf{X}) &\propto p(\boldsymbol{\mu}) \prod_{n=1}^{N} p(\mathbf{x}_n | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0) - \frac{1}{2}\sum_{n=1}^{N}(\mathbf{x}_n - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \boldsymbol{\mu})\right\}
\end{aligned}
$$

We rearrange the exponential term into the following form based on the Lemma.

$$
\begin{aligned}
&-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0) - \frac{1}{2}\sum_{n=1}^{N}(\mathbf{x}_n - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \boldsymbol{\mu}) \\
&= -\frac{1}{2}\boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{\Sigma}_0^{-1} + N\boldsymbol{\Sigma}^{-1})\boldsymbol{\mu} + \boldsymbol{\mu}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}^{-1}\sum_{n=1}^{N}\mathbf{x}_n\right) + \text{const.}
\end{aligned}
$$

In this form, we obtain the parameters of the posterior distribution:

$$
\begin{aligned}
\boldsymbol{\mu}_{\text{post}} &= \left(\boldsymbol{\Sigma}_0^{-1} + N\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}^{-1}\sum_{n=1}^{N}\mathbf{x}_n\right) \\
\boldsymbol{\Sigma}_{\text{post}} &= \left(\boldsymbol{\Sigma}_0^{-1} + N\boldsymbol{\Sigma}^{-1}\right)^{-1}
\end{aligned}
$$

The posterior distribution

$$
p(\boldsymbol{\mu}|\mathbf{X}) = \mathcal{N}\left(\mathbf{x}|\boldsymbol{\mu}_{\text{post}}, \boldsymbol{\Sigma}_{\text{post}}\right).
$$