Kriging distribution under HSGP proof

Christine Shen

March 21, 2026

We claim that under HSGP, the kriging distribution $\boldsymbol{\theta}^* \mid (\boldsymbol{\theta}, \Omega) \stackrel{d}{=} \Phi^* \mathbf{S}^{1/2}\mathbf{b}$ for $\mathbf{b} \sim N_m(\mathbf{0}, \mathbf{I})$.

Proof. We first highlight a few facts that we will use for this proof:

1. By property of multivariate normal, for any jointly normal, and possibly singular normal random vectors

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}\right),$$

   the conditional distribution $\mathbf{Y} \mid \mathbf{X} = \mathbf{x} \sim N(\Sigma_{yx}\Sigma_{xx}^{\dagger}\mathbf{x}, \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{\dagger\top}\Sigma_{xy})$[1], where $\mathbf{A}^\dagger$ denotes a generalized inverse of matrix $\mathbf{A}$ such that $\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$.

2. By the design of HSGP, if $m \leq n$, Φ always has full column rank.

3. By the design of HSGP, $\mathbf{S}$ is always invertible.

By fact 1, under HSGP approximation, $\boldsymbol{\theta}^* \mid (\boldsymbol{\theta}, \Omega) \sim N_q(\mathbb{E}^{HS}_{\boldsymbol{\theta}^*}, \mathbb{V}^{HS}_{\boldsymbol{\theta}^*})$,

$$\begin{aligned}
\mathbb{E}^{HS}_{\boldsymbol{\theta}^*} &= (\Phi^*\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^\dagger\boldsymbol{\theta} \\
\mathbb{V}^{HS}_{\boldsymbol{\theta}^*} &= (\Phi^*\mathbf{S}\Phi^{*\top}) - (\Phi^*\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^{\dagger\top}(\Phi\mathbf{S}\Phi^{*\top}).
\end{aligned}$$

Now we show that if $m \leq n$, $\mathbb{V}^{HS}_{\boldsymbol{\theta}^*} \equiv \mathbf{0}$, and the kriging distribution is degenerate with $\boldsymbol{\theta}^* \mid (\boldsymbol{\theta}, \Omega) = \mathbb{E}^{HS}_{\boldsymbol{\theta}^*}$. This is equivalent to showing

$$(\Phi^*\mathbf{S}\Phi^{*\top}) = (\Phi^*\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^{\dagger\top}(\Phi\mathbf{S}\Phi^{*\top}).$$

It's also equivalent to showing

$$\mathbf{S} = \mathbf{S}\Phi^\top(\Phi\mathbf{S}\Phi^\top)^{\dagger\top}\Phi\mathbf{S}.$$

By fact 2, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$,

$$\Phi\mathbf{x} = \Phi\mathbf{y} \implies \mathbf{x} = \mathbf{y}, \quad \mathbf{x}^\top\Phi^\top = \mathbf{y}^\top\Phi^\top \implies \mathbf{x} = \mathbf{y}.$$

Consequently for any matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times k}$, for any k,

$$\Phi\mathbf{A} = \Phi\mathbf{B} \implies \mathbf{A} = \mathbf{B} \quad \mathbf{A}^\top\Phi^\top = \mathbf{B}^\top\Phi^\top \implies \mathbf{A} = \mathbf{B}. \tag{1}$$

[1] You can find this result on Wiki for multivariate normal distribution, under the section for conditional distribution. I'd like to highlight that if we don't insist on using a symmetric generalized inverse matrix, then there needs to be a "transpose" in the expression for the covariance matrix, i.e., it should be $\Sigma_{xx}^{\dagger\top}$ instead of $\Sigma_{xx}^{\dagger}$. The result on Wiki page is actually using $\Sigma_{xx}^{\dagger}$. This is because, the source for the Wiki results is using a specific form of generalized inverse which is symmetric.

By the definition of generalized inverse, we have

$$(\Phi\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^\dagger(\Phi\mathbf{S}\Phi^\top) = (\Phi\mathbf{S}\Phi^\top). \tag{2}$$

Taking transpose of both sides of the equation, we have

$$(\Phi\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^{\dagger\top}(\Phi\mathbf{S}\Phi^\top) = (\Phi\mathbf{S}\Phi^\top).$$

Now applying equation (1), we have $\mathbf{S}\Phi^\top(\Phi\mathbf{S}\Phi^\top)^{\dagger\top}\Phi\mathbf{S} = \mathbf{S}$, which is exactly what's needed to show $\mathbb{V}^{HS}_{\boldsymbol{\theta}^*} \equiv \mathbf{0}$ as argued above. Applying the same arguments to equation (2) gives

$$\mathbf{S}\Phi^\top(\Phi\mathbf{S}\Phi^\top)^\dagger\Phi\mathbf{S} = \mathbf{S}. \tag{3}$$

Therefore under the reparameterized model, the kriging distribution is

$$\begin{aligned}
\boldsymbol{\theta}^* \mid (\boldsymbol{\theta}, \Omega) &= (\Phi^*\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^\dagger\boldsymbol{\theta} \\
&= (\Phi^*\mathbf{S}\Phi^\top)(\Phi\mathbf{S}\Phi^\top)^\dagger(\Phi\mathbf{S}^{1/2}\mathbf{b}) \\
&= \Phi^*(\mathbf{S}\Phi^\top(\Phi\mathbf{S}\Phi^\top)^\dagger\Phi\mathbf{S})\mathbf{S}^{-1/2}\mathbf{b} \quad \text{by fact 3} \\
&= \Phi^*\mathbf{S}^{1/2}\mathbf{b}. \quad \text{by equation (3)}
\end{aligned}$$

This completes the proof for $m \leq n$ cases.

Next we prove for situation where $m > n$. Note in this case $\mathbf{S}^{1/2}\Phi^\top \in \mathbb{R}^{m\times n}$ is a tall matrix. Let $\mathbf{S}^{1/2}\Phi^\top = \mathbf{QR}$ be its QR decomposition where $\mathbf{R} \in \mathbb{R}^{n\times n}$ is an invertible upper triangular matrix, $\mathbf{Q} \in \mathbb{R}^{m\times n}$ has orthonormal columns such that $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$ and $\mathbf{QQ}^\top$ is an orthogonal projection matrix. Now that $m > n$, we can replace the generalized inverse in the kriging distribution with proper inverse. Then plugging $\mathbf{S}^{1/2}\Phi^\top = \mathbf{QR}$ in the kriging distribution under the reparameterized model gives:

$$\begin{aligned}
\mathbb{E}^{HS}_{\boldsymbol{\theta}^*} &= \Phi^*\mathbf{S}^{1/2}(\mathbf{QR})(\mathbf{R}^\top\mathbf{Q}^\top\mathbf{QR})^{-1}(\mathbf{QR})^\top\mathbf{b} \\
&= \Phi^*\mathbf{S}^{1/2}\mathbf{QQ}^\top\mathbf{b} \\
\mathbb{V}^{HS}_{\boldsymbol{\theta}^*} &= (\Phi^*\mathbf{S}\Phi^{*\top}) - \Phi^*\mathbf{S}^{1/2}(\mathbf{QR})(\mathbf{R}^\top\mathbf{Q}^\top\mathbf{QR})^{-1}(\mathbf{QR})^\top\mathbf{S}^{1/2}\Phi^{*\top} \\
&= (\Phi^*\mathbf{S}\Phi^{*\top}) - \Phi^*\mathbf{S}^{1/2}\mathbf{QQ}^\top\mathbf{S}^{1/2}\Phi^{*\top} \\
&= \Phi^*\mathbf{S}^{1/2}(I - \mathbf{QQ}^\top)\mathbf{S}^{1/2}\Phi^{*\top} \\
&= \Phi^*\mathbf{S}^{1/2}(I - \mathbf{QQ}^\top)^2\mathbf{S}^{1/2}\Phi^{*\top}.
\end{aligned}$$

Therefore, to sample from this distribution, we can generate $\mathbf{b}^* \sim N_m(0, \mathbf{I}) \perp \mathbf{b}$, and then use $\Phi^*\mathbf{S}^{1/2}(\mathbf{I} - \mathbf{QQ}^\top)\mathbf{b}^* + \Phi^*\mathbf{S}^{1/2}\mathbf{QQ}^\top\mathbf{b}$. However, it has the same distribution as $\Phi^*\mathbf{S}^{1/2}\mathbf{b}$ because the covariance between $(\mathbf{I} - \mathbf{QQ}^\top)\mathbf{b}$ and $\mathbf{QQ}^\top\mathbf{b}$ is zero. This completes the proof.

□