

Problem class sheet 3

Lecturer & author: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

Exercise 1. (★★)

Inventory of useful formulas.

[Normal distr. conditioning] Let $x_1 \in \mathbb{R}^{d_1}$, and $x_2 \in \mathbb{R}^{d_2}$. If

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N_{d_1+d_2} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_1 & \Sigma_{21}^\top \\ \Sigma_{21} & \Sigma_2 \end{bmatrix} \right)$$

then it is

$$x_2 | x_1 \sim N_{d_2} (\mu_{2|1}, \Sigma_{2|1})$$

where

$$\mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_1^{-1}(x_1 - \mu_1) \quad \text{and} \quad \Sigma_{2|1} = \Sigma_2 - \Sigma_{21}\Sigma_1^{-1}\Sigma_{21}^\top$$

Consider the Bayesian Kriging from your lecture notes:

$$Z(s) = Y(s) + \varepsilon(s), \quad s \in \mathcal{S}$$

where

$$\varepsilon(\cdot) \sim \text{GP}(0, c_\varepsilon(\cdot, \cdot | \tau))$$

with $c_\varepsilon(s, s' | \tau) = \tau^2 1_{\{0\}}(\|s - s'\|)$ and

$$Y(\cdot) | \beta, \theta \sim \text{GP}(\mu(\cdot | \beta), c_Y(\cdot, \cdot | \sigma^2, \phi))$$

with mean function $\mu(\cdot | \beta)$ (to be specified later) labeled by unknown parameter β and covariance function $c_Y(\cdot, \cdot | \sigma^2, \phi)$.

Assume there is available a dataset $\{(s_i, Z_i)\}_{i=1}^n$ where $Z_i = Z(s_i)$ is a realization of a stochastic process (Z_s) .

- (1) Write the hierarchical spatial model $Z(\cdot) | Y(\cdot), \beta, \varphi$ and $Y(\cdot) | \beta, \varphi$ where $\varphi = (\sigma^2, \phi, \tau)^\top$.
- (2) Write the marginal process $Z(\cdot) | \beta, \varphi$ where $\varphi = (\sigma^2, \phi, \tau)^\top$, its mean function denoted as $\mu(\cdot | \cdot)$, and its covariance function denoted as $c(\cdot | \cdot)$.

Hint:: Let Y and X be independent random variables with $X \sim N(\mu_X, \Sigma_X)$, $Y \sim N(\mu_Y, \Sigma_Y)$. Let A and B be fixed matrices. Let c be a fixed vector. Then

$$AX + BY + c \sim N(A\mu_X + B\mu_Y + c, A\Sigma_X A^\top + B\Sigma_Y B^\top)$$

(3) Compute the predictive process $Z(\cdot) | Z, \beta, \varphi$ as

$$Z(\cdot) | Z, \beta, \varphi \sim \text{GP}(\mu_1(\cdot | \beta, \varphi), c_1(\cdot, \cdot | \varphi))$$

with

$$\begin{aligned} c_1(s, s' | \varphi) &= c(s, s | \varphi) + (C(S, s | \varphi))^\top (C(S, S | \varphi))^{-1} C(S, s' | \varphi) \\ \mu_1(s | \beta, \varphi) &= \mu(s | \beta) - (C(S, s | \varphi))^\top (C(S, S | \varphi))^{-1} (\mu(S | \beta) - Z) \end{aligned}$$

Hint: See the Conditional Normal formula above.

(4) Assume $\mu(s | \beta) = \psi(s)^\top \beta$. Consider a conjugate prior $\beta \sim \text{N}(b, B)$ on β where $B > 0$.

- (a) Write down the Bayesian statistical model involving layers $[Z | \beta, \varphi]$, and $[\beta | \varphi]$.
- (b) Compute the posterior distribution as

$$\beta | Z, \varphi \sim \text{N}(b_n(\varphi), B_n(\varphi))$$

with

$$\begin{aligned} B_n(\varphi) &= (B^{-1} + \Psi^\top (C(S, S | \varphi))^{-1} \Psi)^{-1} \\ b_n(\varphi) &= B_n(\varphi) (B^{-1} b + \Psi^\top (C(S, S | \varphi))^{-1} Z) \end{aligned}$$

where $C(S, S | \varphi)$ is a matrix with $[C(S, S | \varphi)]_{i,j} = c(s_i, s_j | \varphi)$.

Hint: Use the following identity

$$\begin{aligned} (y - \Phi\beta)^\top \Sigma^{-1} (y - \Phi\beta) + (\beta - \mu)^\top V^{-1} (\beta - \mu) &= (\beta - \mu^*)^\top (V^*)^{-1} (\beta - \mu^*) + S^*; \\ V^* &= (V^{-1} + \Phi^\top \Sigma^{-1} \Phi)^{-1}; \quad \mu^* = V^* (V^{-1} \mu + \Phi^\top \Sigma^{-1} y) \\ S^* &= \mu^\top V^{-1} \mu - (\mu^*)^\top (V^*)^{-1} (\mu^*) + y^\top \Sigma^{-1} y; \end{aligned}$$

(c) Compute the (posterior) predictive process $Z(\cdot) | Z, \varphi$ given the data Z and given the parameters φ as

$$Z(\cdot) | Z, \varphi \sim \text{GP}(\mu_2(\cdot | \varphi), c_2(\cdot, \cdot | \varphi))$$

with

$$\begin{aligned} \mu_2(s | \varphi) &= (\psi(s) - \Psi^\top C^{-1} C(s))^\top (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} B^{-1} b \\ &\quad + \left[(C(s))^\top + (\psi(s) - \Psi^\top C^{-1} C(s))^\top (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} \Psi \right] C^{-1} Z \\ c_2(s, s' | \varphi) &= c(s, s' | \varphi) - (C(s))^\top C^{-1} C(s') \\ &\quad + (\psi(s) - \Psi^\top C^{-1} C(s))^\top (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} (\psi(s') - \Psi^\top C^{-1} C(s')) \end{aligned}$$

with column vector $C(s) := (c(s, s_1 | \varphi), \dots, c(s, s_n | \varphi))^\top$, and matrix $C := C(S, S | \varphi)$.

(d) Compute the marginal likelihood $\Pr(Z|\varphi)$ in the form

$$\Pr(Z|\sigma^2, \varphi) = N\left(Z|\Psi b, \left(C^{-1} - C^{-1}\Psi(B^{-1} + \Psi^\top B^{-1}\Psi)^{-1}\Psi^\top C^{-1}\right)^{-1}\right)$$

where Ψ is a matrix with $[\Psi]_{i,j} = \psi_j(s_i)$, and R is a matrix with $[C]_{i,j} = c(s_i, s_j|\varphi)$.

Hint-2:: It is

$$\int N(Z|\Psi\beta, C) N(\beta|b, B) d\beta = N(Z|\Psi b, C + \Psi B \Psi^\top)$$

Hint 3:: [Woodbury matrix identity]

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$$

(5) Consider non-informative prior $\Pr(\beta) \propto 1$ for β by specifying $b \rightarrow 0$ and letting $B^{-1} \rightarrow 0$. Argue whether such a prior can be used. Recompute the (asymptotic) quantities $\Pr(Z|\varphi)$, $[Z(\cdot)|Z, \varphi]$ under this new prior in the limit.

Solution.

(1) The hierarchical model is

$$\begin{aligned} Z(\cdot)|Y(\cdot), \tau &\sim \text{GP}(Y(\cdot), c_\varepsilon(\cdot, \cdot|\sigma^2, \xi)) \\ Y(\cdot)|\beta &\sim \text{GP}(\mu(\cdot|\beta), c_Y(\cdot, \cdot|\sigma^2, \phi)) \end{aligned}$$

(2) We use the additive property of the Gaussian distribution (in Hint) it is

$$Z(\cdot)|\beta, \varphi \sim \text{GP}(\mu(\cdot|\beta), c(\cdot, \cdot|\varphi))$$

where

$$c(s, s'|\varphi) = c_Y(s, s'|\sigma^2, \phi) + c_\varepsilon(s, s'|\sigma^2, \xi)$$

(3) Assume a vector of “unseen” sites $S_* = (s_{*,1}, \dots, s_{*,q})^\top$ for any $q \in \mathbb{N}_0$. Let convenient notation $Z := Z(S)$, and $Z_* := Z(S_*)$. The joint marginal distribution of $(Z_*, Z)^\top$ given $\beta, \varphi = (\sigma^2, \phi, \tau)^\top$ is

$$\begin{pmatrix} Z_* \\ Z \end{pmatrix} | \beta, \varphi \sim N\left(\begin{pmatrix} \mu(S_*; \beta) \\ \mu(S; \beta) \end{pmatrix}, \begin{pmatrix} C(S_*, S_*|\varphi) & (C(S_*, S|\varphi))^\top \\ C(S_*, S|\varphi) & C(S, S|\varphi) \end{pmatrix}\right)$$

by using convenient notation $[C(S_*, S|\varphi)]_{i,j} = c(s_{*,i}, s_j|\varphi)$ and $[\mu(S; \beta)]_i = \mu(s_i; \beta)$. By conditioning the Normal distribution (see Hint), I get

$$Z_*|Z, \beta, \varphi \sim N(\mu_*(S_*|\beta, \varphi), C_*(S_*, S_*|\varphi))$$

where

$$\begin{aligned} C_1(S_*, S_*|\varphi) &= C(S_*, S_*|\varphi) + (C(S, S_*|\varphi))^\top (C(S, S|\varphi))^{-1} C(S, S_*|\varphi) \\ \mu_1(S_*|\beta, \varphi) &= \mu(S_*|\beta) - (C(S, S_*|\varphi))^\top (C(S, S|\varphi))^{-1} (\mu(S|\beta) - Z) \end{aligned}$$

As it is for any length of of any vector S_* , then it is a Gaussian process

$$Z(\cdot) | Z, \beta, \varphi \sim \text{GP}(\mu_1(\cdot|\beta, \varphi), c_1(\cdot, \cdot|\varphi))$$

with

$$\begin{aligned} c_1(s, s'|\varphi) &= c(s, s'|\varphi) + (C(S, s|\varphi))^\top (C(S, S|\varphi))^{-1} C(S, s'|\varphi) \\ \mu_1(s|\beta, \varphi) &= \mu(s|\beta) - (C(S, s|\varphi))^\top (C(S, S|\varphi))^{-1} (\mu(S|\beta) - Z) \end{aligned}$$

(4)

(a) The Bayesian model is

$$(0.1) \quad \begin{cases} Z|\beta, \varphi \sim \text{N}(\Psi\beta, C(S, S|\varphi)) \\ \beta \sim \text{N}(b, B) \end{cases}$$

(b) Let $C := C(S, S|\varphi)$. The posterior distribution (by using Bayes theorem) is

$$\begin{aligned} \Pr(\beta|Z, \varphi) &\propto \Pr(Z|\beta, \varphi) \Pr(\beta|\varphi) \\ &= \text{N}(Z|\Psi\beta, C) \text{N}(\beta|b, B) \\ &\propto \exp\left(-\frac{1}{2}(Z - \Psi\beta)^\top C^{-1}(Z - \Psi\beta)\right) \exp\left(-\frac{1}{2}(\beta - b)^\top B^{-1}(\beta - b)\right) \\ &= \exp\left(-\frac{1}{2}\left[(Z - \Psi\beta)^\top C^{-1}(Z - \Psi\beta) + (\beta - b)^\top B^{-1}(\beta - b)\right]\right) \end{aligned}$$

By using the Hint I have

$$(Z - \Psi\beta)^\top C^{-1}(Z - \Psi\beta) + (\beta - b)^\top B^{-1}(\beta - b) = (\beta - b_n)^\top (B_n)^{-1}(\beta - b_n) + R_n$$

where by denoting $B_n := B_n(\varphi)$, and $b_n := b_n(\varphi)$ I get

$$\begin{aligned} B_n &= (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} \\ b_n &= B_n (B^{-1} b + \Psi^\top C^{-1} Z) \end{aligned}$$

and R_n is a “constant” quantity that does not contain any β . Hence

$$\begin{aligned} \Pr(\beta|Z, \varphi) &\propto \exp\left(-\frac{1}{2}(\beta - b_n)^\top (B_n)^{-1}(\beta - b_n) - \frac{1}{2}R_n\right) \\ &\propto \exp\left(-\frac{1}{2}(\beta - b_n)^\top (B_n)^{-1}(\beta - b_n)\right) \end{aligned}$$

Well, from the above, I recognize the kernel of the Multivariate Normal distribution, as

$$\beta|Z, \varphi \sim N(b_n(\varphi), B_n(\varphi))$$

- (c) Assume a vector of “unseen” sites $S_* = (s_{*,1}, \dots, s_{*,q})^\top$ for any $q \in \mathbb{N} - \{0\}$. Let convenient notation $Z := Z(S)$, and $Z_* := Z(S_*)$. I have already computed

$$\Pr(Z_*|Z, \beta, \varphi) = N(Z_*|\mu_1(S_*|\beta, \varphi), C_1(S_*, S_*|\varphi))$$

from the previous part. It is

$$\begin{aligned} \Pr(Z_*|Z, \varphi) &= \int \Pr(Z_*|Z, \beta, \varphi) \Pr(\beta|Z, \varphi) d\beta \\ &= \int N(Z_*|\mu_1(S_*|\beta, \varphi), C_1(S_*, S_*|\varphi)) N(\beta|b_n, B_n) d\beta \end{aligned}$$

Denote $\Psi_* = \Psi(S_*)$, $C_* = C(S_*, S|\varphi)$, and $C_{**} = C(S_*, S_*|\varphi)$. Notice that

$$\begin{aligned} \mu_1(S_*) &= \Psi_*\beta - C_*C^{-1}(\Psi\beta - Z) \\ &= [\Psi_* - C_*C^{-1}\Psi]\beta + C_*C^{-1}Z \end{aligned}$$

Hence, for given/fixed Z, φ , it is

$$Z_* = C_*C^{-1}Z + [\Psi_* - C_*C^{-1}\Psi]\beta + \zeta, \quad \zeta \sim N(0, C_1(S_*, S_*))$$

Hence, because $\beta \sim N(b_n, B_n)$, and because $Z_*|Z, \varphi$ is a linear combination of the Normally distributed random vector $\beta \sim N(b_n, B_n)$, $Z_*|Z, \varphi$ follows a Normal distribution, with mean

$$\begin{aligned} \mu_2(S_*) &= E_{\beta \sim N(b_n, B_n)}(Z_*|\mu_1(S_*), C_1(S_*, S_*)) \\ &= (\Psi_* - C_*C^{-1}\Psi) E_{\beta \sim N(b_n, B_n)}(\beta) + C_*C^{-1}Z \\ &= (\Psi_* - C_*C^{-1}\Psi)b_n + C_*C^{-1}Z \\ &= (\Psi_* - C_*C^{-1}\Psi)(B^{-1} + \Psi^\top C^{-1}\Psi)^{-1}(B^{-1}b + \Psi^\top C^{-1}Z) + C_*C^{-1}Z \\ &= (\Psi_* - C_*C^{-1}\Psi)(B^{-1} + \Psi^\top C^{-1}\Psi)^{-1}B^{-1}b \\ &\quad + [(\Psi_* - C_*C^{-1}\Psi)(B^{-1} + \Psi^\top C^{-1}\Psi)^{-1}\Psi^\top + C_*]C^{-1}Z \end{aligned}$$

and with covariance matrix

$$\begin{aligned}
C_2(S_*, S_*) &= \text{Var}_{\beta \sim N(b_n, B_n)}(Z_* | \mu_1(S_*), C_1(S_*, S_*)) \\
&= \text{Var}_{\beta \sim N(b_n, B_n)}([\Psi_* - C_* C^{-1} \Psi] \beta) + \text{Var}_{\zeta \sim N(0, C_1(S_*, S_*))}(\zeta) \\
&= [\Psi_* - C_* C^{-1} \Psi] B_n [\Psi_* - C_* C^{-1} \Psi]^\top + C_1(S_*, S_*) \\
&= [\Psi_* - C_* C^{-1} \Psi] (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} [\Psi_* - C_* C^{-1} \Psi]^\top \\
&\quad + C_{**} + C_* C^{-1} (C_*)^\top
\end{aligned}$$

Recall that $C(s) = (c(s_1, s|\varphi), \dots, c(s_n, s|\varphi))^\top$ is a column vector.

Since this is for any vector S_* of any length, then

$$Z(\cdot) | Z, \varphi \sim \text{GP}(\mu_2(\cdot|\varphi), c_2(\cdot, \cdot|\varphi))$$

with mean function at s

$$\begin{aligned}
\mu_2(s|\varphi) &= \left(\psi(s) - (C(s))^\top C^{-1} \Psi \right) (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} B^{-1} b \\
&\quad + \left[\left(\psi(s) - (C(s))^\top C^{-1} \Psi \right) (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} \Psi^\top + (C(s))^\top \right] C^{-1} Z
\end{aligned}$$

and covariance function as s, s'

$$\begin{aligned}
c_2(s, s'|\varphi) &= \left[\psi(s) - (C(s))^\top C^{-1} \Psi \right] (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} \left[\psi(s') - (C(s'))^\top C^{-1} \Psi \right]^\top \\
&\quad + c(s, s'|\varphi) + (C(s))^\top C^{-1} C(s')
\end{aligned}$$

Recall that $C(s) = (c(s_1, s|\varphi), \dots, c(s_n, s|\varphi))^\top$ is a column vector.

(d) It is, from Hint-2,

$$\begin{aligned}
\Pr(Z|\varphi) &= \int \Pr(Z|\beta, \varphi) \Pr(\beta) d\beta \\
&= \int N(Z|\Psi\beta, C(S, S|\varphi)) N(\beta|b, B) d\beta \\
&= \int N(Z|\Psi\beta, C(S, S|\varphi)) N(\Psi\beta|\Psi b, \Psi B \Psi^\top) d\beta \\
&= N(Z|\Psi b, C(S, S|\varphi) + \Psi B \Psi^\top)
\end{aligned}$$

By letting $C := C(S, S|\varphi)$ and using the Hint I get

$$\left((C + \Psi B \Psi^\top)^{-1} \right)^{-1} = \left(C^{-1} - C^{-1} \Psi (B^{-1} + \Psi^\top C^{-1} \Psi)^{-1} \Psi^\top C^{-1} \right)^{-1}$$

(5) Denote $C = C(S, S|\varphi)$. It is

$$\begin{aligned} \lim_{\substack{B^{-1} \rightarrow 0 \\ b \rightarrow 0}} \Pr(Z|\varphi) &= \lim_{\substack{B^{-1} \rightarrow 0 \\ b \rightarrow 0}} \mathcal{N}(Z|\Psi b, C + \Psi B \Psi^\top) \\ &\propto \mathcal{N}\left(Z|0, \left(C^{-1} - C^{-1}\Psi(\Psi^\top C^{-1}\Psi)^{-1}\Psi^\top C^{-1}\right)^{-1}\right) \\ &< \infty \end{aligned}$$

namely the bottom part of the fraction of the posterior of $\beta|Z, \varphi$ is bounded (finite); this implies that the posterior is proper. The posterior of $\beta|Z, \varphi$ has density such as

$$\Pr(\beta|Z, \varphi) \propto \exp\left(-\frac{1}{2}(\beta - b_n)^\top B_n^{-1}(\beta - b_n)\right)$$

then by computing the limit

$$\begin{aligned} \lim_{\substack{B^{-1} \rightarrow 0 \\ b \rightarrow 0}} \exp\left(-\frac{1}{2}(\beta - b_n)^\top B_n^{-1}(\beta - b_n)\right) &= \\ \exp\left(-\frac{1}{2}\left(\beta - (\Psi^\top C^{-1}\Psi)^{-1}\Psi^\top C^{-1}Z\right)^\top (\Psi^\top C^{-1}\Psi) \left(\beta - (\Psi^\top C^{-1}\Psi)^{-1}\Psi^\top C^{-1}Z\right)\right) \end{aligned}$$

Hence the limiting case is

$$\beta|Z, \varphi \stackrel{\text{approx}}{\sim} \mathcal{N}\left((\Psi^\top C^{-1}\Psi)^{-1}\Psi^\top C^{-1}Z, (\Psi^\top C^{-1}\Psi)^{-1}\right)$$

Hence the predictive process becomes

$$Z(\cdot)|Z, \varphi \stackrel{\text{approx}}{\sim} \text{GP}(\mu_3(\cdot|\varphi), c_3(\cdot, \cdot|\varphi))$$

$$\begin{aligned} \mu_3(s|\varphi) &= \left[\left(\psi(s) - (C(s))^\top C^{-1}\Psi \right) (\Psi^\top C^{-1}\Psi)^{-1} \Psi^\top + (C(s))^\top \right] C^{-1}Z \\ c_3(s, s'|\varphi) &= \left(\psi(s) - (C(s))^\top C^{-1}\Psi \right) (\Psi^\top C^{-1}\Psi)^{-1} \left(\psi(s') - (C(s'))^\top C^{-1}\Psi \right)^\top \\ &\quad + c(s, s'|\varphi) + (C(s))^\top C^{-1}C(s') \end{aligned}$$

Exercise 2. (★) Show that the extension variance $\sigma_E^2(v, V)$ of a small volume v to a larger volume V is obtained by

$$\sigma_E^2(v, V) = 2\bar{\gamma}(v, V) - \bar{\gamma}(v, v) - \bar{\gamma}(V, V)$$

where

$$\bar{\gamma}(v, V) = \frac{1}{|v||V|} \int_{s \in v} \int_{s' \in V} \gamma(s - s') \, ds ds'$$

Solution. Essentially I need to show that that

$$\begin{aligned}\text{Var}(Z(A) - Z(B)) &= \frac{1}{|A||B|} \int_{x \in A} \int_{y \in B} \gamma(x - y) \, dx dy \\ &\quad - \frac{1}{|A||A|} \int_{x \in A} \int_{y \in A} \gamma(x - y) \, dx dy \\ &\quad - \frac{1}{|B||B|} \int_{x \in B} \int_{y \in B} \gamma(x - y) \, dx dy\end{aligned}$$

where I use A, B instead of v, V and x, y instead of s, s' for clarity on notation.

It is

$$\begin{aligned}\text{Var}(Z(A) - Z(B)) &= \text{Cov}(Z(A) - Z(B), Z(A) - Z(B)) \\ &= \text{Cov}(Z(A), Z(A)) + \text{Cov}(Z(B), Z(B)) - 2\text{Cov}(Z(A), Z(B)) \\ &= \frac{1}{|A||A|} \int_{x \in A} \int_{y \in A} \text{Cov}(Z(x), Z(y)) \, dx dy \\ &\quad + \frac{1}{|B||B|} \int_{x \in B} \int_{y \in B} \text{Cov}(Z(x), Z(y)) \, dx dy \\ &\quad - 2 \frac{1}{|A||B|} \int_{x \in A} \int_{y \in B} \text{Cov}(Z(x), Z(y)) \, dx dy\end{aligned}$$

OK, now I need to write all these Cov as γ ; I know that

$$\begin{aligned}\gamma(x - y) &= \frac{1}{2} \text{Var}(Z(x) - Z(y)) \\ &= \frac{1}{2} \text{Var}(Z(x)) + \frac{1}{2} \text{Var}(Z(y)) - \text{Cov}(Z(x), Z(y))\end{aligned}$$

that is

$$\text{Cov}(Z(x), Z(y)) = \frac{1}{2} \text{Var}(Z(x)) + \frac{1}{2} \text{Var}(Z(y)) - \gamma(x - y)$$

Now I'll gonna put all these in the quantity of interest, one by one

$$\begin{aligned}\frac{1}{|A||A|} \int_{x \in A} \int_{y \in A} \text{Cov}(Z(x), Z(y)) \, dx dy &= \frac{1}{|A||A|} \int_{x \in A} \int_{y \in A} \frac{1}{2} \text{Var}(Z(x)) \, dx dy \\ &\quad + \frac{1}{|A||A|} \int_{x \in A} \int_{y \in A} \frac{1}{2} \text{Var}(Z(y)) \, dx dy \\ &\quad - \frac{1}{|A||A|} \int_{x \in A} \int_{y \in A} \gamma(x - y) \, dx dy \\ &= \frac{1}{|A|} \int_{x \in A} \text{Var}(Z(x)) \, dx \\ &\quad - \frac{1}{|A|^2} \int_{x \in A} \int_{y \in A} \gamma(x - y) \, dx dy\end{aligned}$$

and by symmetry

$$\begin{aligned} \frac{1}{|B| |B|} \int_{x \in B} \int_{y \in B} \text{Cov}(Z(x), Z(y)) \, dx dy &= \frac{1}{|B|} \int_{x \in B} \text{Var}(Z(x)) \, dx \\ &\quad - \frac{1}{|B|^2} \int_{x \in B} \int_{y \in B} \gamma(x-y) \, dx dy \end{aligned}$$

and finally,

$$\begin{aligned} \frac{1}{|A| |B|} \int_{x \in A} \int_{y \in B} \text{Cov}(Z(x), Z(y)) \, dx dy &= \frac{1}{|A| |B|} \int_{x \in A} \int_{y \in B} \frac{1}{2} \text{Var}(Z(x)) \, dx dy \\ &\quad + \frac{1}{|A| |B|} \int_{x \in A} \int_{y \in B} \frac{1}{2} \text{Var}(Z(y)) \, dx dy \\ &\quad - \frac{1}{|A| |B|} \int_{x \in A} \int_{y \in B} \gamma(x-y) \, dx dy \\ &= \frac{1}{2} \frac{1}{|A|} \int_{x \in A} \text{Var}(Z(x)) \, dx \\ &\quad + \frac{1}{2} \frac{1}{|B|} \int_{x \in B} \text{Var}(Z(x)) \, dx \\ &\quad - \frac{1}{|A| |B|} \int_{x \in A} \int_{y \in B} \gamma(x-y) \, dx dy \end{aligned}$$

Putting all these together, we get the result.
