Notes - A GP model for shear fields

Karen Y. Ng

December 19, 2014

Our Gaussian process "model" for the projected lensing potential

$$\psi \sim N(0, \Sigma(\vec{x}, \vec{y})) \tag{1}$$

which is a scalar field evaluated at the positions $x_i = (\vec{x_1}, \vec{x_2})_i$ or $y_i = (\vec{y_1}, \vec{y_2})_i$ where we have data points. As usual, first column, x_1 or y_1 is the first spatial dimension, x_2 or y_2 is the second one, the i-th row correspond to spatial coordinates of the i-th data point.
$\vec{x}$ and $\vec{y}$ are the same but we call them different names for denoting their location in the covariance matrix . . . .

For inferring the convergence and shear, we need the 2nd spatial derivatives. The subscripts in these WL equations correspond to the spatial coordinates x_1, x_2 NOT the observation numbers i.e. i, j = 1, 2, . . . , n observations

$$\begin{aligned}\kappa &= \frac{1}{2} tr(\psi_{,ij}) \\ &= \frac{1}{2}(\psi_{,11} + \psi_{,22}) \\ &= \frac{1}{2}\left(\frac{\partial^2\psi}{\partial x_1^2} + \frac{\partial^2\psi}{\partial x_2^2}\right)\end{aligned}$$

$$\begin{aligned}\gamma_1 &= \frac{1}{2}(\psi_{,11} - \psi_{,22}) \\ &= \frac{1}{2}\left(\frac{\partial^2\psi}{\partial x_1^2} - \frac{\partial^2\psi}{\partial x_2^2}\right)\end{aligned}$$

$$\begin{aligned}\gamma_2 &= \frac{1}{2}(\psi_{,12} + \psi_{,21}) \\ &= \frac{1}{2}\left(\frac{\partial^2\psi}{\partial x_1 \partial x_2} + \frac{\partial^2\psi}{\partial x_2 \partial x_1}\right)\end{aligned}$$

Covariances of the required functions

Note that ψ, κ and γ are scalar fields. However, we are evaluating them at the locations of the data points $(x_1, x_2)_i$, therefore, when we are writing down the shorthand for the m, n subscripts below, we mean, we first take the spatial derivatives of those scalar field(s) with respect to x_1 or x_2, then evaluate them at the $m-th$ or $n-th$ position $(x_1, x_2)_m$. The spatial derivatives are represented as follows:

$$\psi_{,1} = \frac{\partial \psi}{\partial x_1}$$

etc. with a comma in the subscript.

Also note expectation and derivative are both linear operators, so we can exchange their positions (and try not to let mathematicians read this and shoot us)

$$\begin{aligned}
&\mathrm{Cov}_{m,n}(\kappa(\vec{x}), \kappa(\vec{y})) \\
&= \mathbb{E}\left[(\kappa - \mathbb{E}[\kappa])|_m (\kappa - \mathbb{E}[\kappa])|_n\right] \\
&= \mathbb{E}\left[\left[\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right)\psi - \mathbb{E}\left[\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right)\psi\right]\right]\Big|_m \left[\frac{1}{2}\left(\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial y_2^2}\right)\psi - \mathbb{E}\left[\frac{1}{2}\left(\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial y_2^2}\right)\psi\right]\right]\Big|_n\right] \\
&= \frac{1}{4}\mathbb{E}\left[\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right)\Big|_m [\psi - \mathbb{E}[\psi]]|_m \left(\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial y_2^2}\right)\Big|_n [\psi - \mathbb{E}[\psi]]|_n\right] \\
&= \frac{1}{4}\left(\left(\frac{\partial^2}{\partial x_1^2}\right)\Big|_m \left(\frac{\partial^2}{\partial y_1^2}\right)\Big|_n + \left(\frac{\partial^2}{\partial x_1^2}\right)\Big|_m \left(\frac{\partial^2}{\partial y_2^2}\right)\Big|_n + \left(\frac{\partial^2}{\partial x_2^2}\right)\Big|_m \left(\frac{\partial^2}{\partial y_1^2}\right)\Big|_n + \left(\frac{\partial^2}{\partial x_2^2}\right)\Big|_m \left(\frac{\partial^2}{\partial y_2^2}\right)\Big|_n\right)\Sigma_{mn} \\
&= \frac{1}{4}\left(\frac{\partial^2}{\partial x_1^2}\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial x_1^2}\frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial x_2^2}\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial x_2^2}\frac{\partial^2}{\partial y_2^2}\right)\Sigma_{mn}
\end{aligned}$$

Now I have dropped the (m,n) subscripts and the following subscripts correspond to the spatial dimensions, the first two subscripts correspond to spatial derivatives w.r.t. x and evaluated for the m-th data points, the last two correspond to spatial derivatives w.r.t. y and evaluated for the n-th data points.

$$\mathrm{Cov}(\kappa(\vec{x}), \kappa(\vec{y})) = \frac{1}{4}\left(\Sigma_{,1111} + \Sigma_{,1122} + \Sigma_{,2211} + \Sigma_{,2222}\right) \tag{2}$$

Similarly,

$$\begin{aligned}
&\mathrm{Cov}_{mn}(\gamma_1(\vec{x}), \gamma_1(\vec{y})) \\
&= \mathbb{E}\left[(\gamma_1 - \mathbb{E}[\gamma_1])|_m (\gamma_1 - \mathbb{E}[\gamma_1])|_n\right] \\
&= \mathbb{E}\left[\left[\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2}\right)\psi - \mathbb{E}\left[\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2}\right)\psi\right]\right]\Big|_m \left[\frac{1}{2}\left(\frac{\partial^2}{\partial y_1^2} - \frac{\partial^2}{\partial y_2^2}\right)\psi - \mathbb{E}\left[\frac{1}{2}\left(\frac{\partial^2}{\partial y_1^2} - \frac{\partial^2}{\partial y_2^2}\right)\psi\right]\right]\Big|_n\right] \\
&= \frac{1}{4}\mathbb{E}\left[\left(\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2}\right)\Big|_i [\psi - \mathbb{E}[\psi]]|_m \left(\frac{\partial^2}{\partial y_1^2} - \frac{\partial^2}{\partial y_2^2}\right)\Big|_j [\psi - \mathbb{E}[\psi]]|_n\right] \\
&= \frac{1}{4}\left(\frac{\partial^2}{\partial x_1^2}\frac{\partial^2}{\partial y_1^2} - \frac{\partial^2}{\partial x_1^2}\frac{\partial^2}{\partial y_2^2} - \frac{\partial^2}{\partial x_2^2}\frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial x_2^2}\frac{\partial^2}{\partial y_2^2}\right)\Sigma_{mn}
\end{aligned}$$

$$\mathrm{Cov}(\gamma_1(\vec{x}), \gamma_1(\vec{y})) = \frac{1}{4}\left(\Sigma_{,1111} - \Sigma_{,1122} - \Sigma_{,2211} + \Sigma_{,2222}\right) \tag{3}$$

And,

$$
\begin{aligned}
\mathrm{Cov}_{mn}(\gamma_2(\vec{x}), \gamma_2(\vec{y})) &= \mathbb{E}\left[(\gamma_2 - \mathbb{E}[\gamma_2])|_m (\gamma_2 - \mathbb{E}[\gamma_2])|_n\right] \\
&= \mathbb{E}\Bigg[\left[\frac{1}{2}\left(\frac{\partial^2}{\partial x_1 \partial x_2} + \frac{\partial^2}{\partial x_2 \partial x_1}\right)\psi - \mathbb{E}\left[\frac{1}{2}\left(\frac{\partial^2}{\partial x_1 \partial x_2} + \frac{\partial^2}{\partial x_2 \partial x_1}\right)\psi\right]\right]\Bigg|_m \\
&\qquad \left[\frac{1}{2}\left(\frac{\partial^2}{\partial y_1 \partial y_2} + \frac{\partial^2}{\partial y_2 \partial y_1}\right)\psi - \mathbb{E}\left[\frac{1}{2}\left(\frac{\partial^2}{\partial y_1 \partial y_2} + \frac{\partial^2}{\partial y_2 \partial y_1}\right)\psi\right]\right]\Bigg|_n\Bigg]
\end{aligned}
$$

$$
\mathrm{Cov}_{mn}(\gamma_2(\vec{x}), \gamma_2(\vec{y})) = \frac{1}{4}\mathbb{E}\left[\left(\frac{\partial^2}{\partial x_1 \partial x_2} + \frac{\partial^2}{\partial x_2 \partial x_1}\right)\Bigg|_m [\psi - \mathbb{E}[\psi]]|_m \left(\frac{\partial^2}{\partial y_1 \partial y_2} + \frac{\partial^2}{\partial y_2 \partial y_1}\right)\Bigg|_n [\psi - \mathbb{E}[\psi]]|_n\right]
$$

$$
\mathrm{Cov}(\gamma_2(\vec{x}), \gamma_2(\vec{y})) = \frac{1}{4}\left(\Sigma_{,1212} + \Sigma_{,1221} + \Sigma_{,2112} + \Sigma_{,2121}\right) \tag{4}
$$

The squared exponential covariance function

$$
\Sigma(r^2; \lambda, \rho) = \lambda^{-1} \exp\left(-\frac{\beta}{2} r^2\right) \tag{5}
$$

where $\beta = -1/4 \ln \rho$, and $0 < \rho < 1$, note Σ is an N $\times$ N matrix and the covariance functions of the derivatives should have the same dimension.

The metric D

$$
r^2 = (\vec{x} - \vec{y})^T D (\vec{x} - \vec{y}) \tag{6}
$$

Since we are working in projected (2D) space, D is a 2 $\times$ 2 matrix. More explicitly, I will use i,j,h,k as subscripts for the spatial dimensions and m, n for the observation number in the GP model:

$$
\begin{aligned}
r^2_{mn} &= (x_{m1} - y_{m1}, x_{n2} - y_{n2}) \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} x_{m1} - y_{n1} \\ x_{m2} - y_{n2} \end{pmatrix} \\
\Sigma_{mn} &= \lambda^{-1} \exp\left(-\frac{\beta}{2} r^2_{mn}\right)
\end{aligned}
$$

An example of r^2 with an Euclidean metric for a pair of data points, $\vec{x}_i$ and $\vec{y}_j$ would be:

$$
r^2_{mn} = (x_{m1} - y_{n1})^2 + (x_{m2} - y_{n2})^2
$$

In the following derivations, it is NOT important to keep the m, n subscripts. We are taking the derivatives w.r.t to the spatial dimensions, so I will drop the m, n subscripts. But keep in mind each forth derivative of Σ with the indices written out should be a scalar, each is an element in the big $m \times n$ $\Sigma_{,x_i x_j y_h y_k}$ matrix.

Summary: basic derivatives of components, assuming diagonal D

Derivation of the derivatives for a non-symmetric / non-diagonal D would give more terms ...

$$
\begin{aligned}
\frac{\partial r^2}{\partial x_i} &= \frac{\partial (x_q - y_q) D_{qr} (x_r - y_r)}{\partial x_i} \\
&= \delta_{iq} D_{qr} (x_r - y_r) + (x_q - y_q) \delta_{iq} D_{qr} \\
&= 2 D_{ii} (x_i - y_i) \\
&= 2[D(\vec{x} - \vec{y})]_i
\end{aligned}
$$

The third line in the above derivation is only true for diagonal D or else there are other summation terms.

$$
r_{,x_i} = 2[D(\vec{x} - \vec{y})]_i \equiv 2X_i \tag{7}
$$

Similarly,

$$
\begin{aligned}
\frac{\partial r^2}{\partial y_h} &= \frac{\partial (x_q - y_q) D_{qr} (x_r - y_r)}{\partial y_h} \\
&= -\delta_{iq} D_{qr} (x_r - y_r) - (x_q - y_q) \delta_{iq} D_{qr} \\
&= -2[D(\vec{x} - \vec{y})]_h
\end{aligned}
$$

$$
r_{,y_h} = -2[D(\vec{x} - \vec{y})]_h \equiv -2X_h \tag{8}
$$

where i = 1, 2

Second derivatives of r^2

$$
r^2_{,x_i x_j} = 2\delta_{ij} D_{ij} \tag{9}
$$

$$
r^2_{,y_h y_k} = 2\delta_{hk} D_{hk} \tag{10}
$$

$$
r^2_{,x_i y_h} = -2\delta_{ih} D_{ih} \tag{11}
$$

$$
X_i = [D(\vec{x} - \vec{y})]_i \tag{12}
$$

$$
X_i, x_j = D\delta_{ij} \tag{13}
$$

$$
X_i, y_h = -D\delta_{ih} \tag{14}
$$

derivatives of the kernel

$$
\Sigma = \lambda^{-1} k
$$

$$
k = \exp\left(\frac{-\beta}{2} r^2\right) \tag{15}
$$

$$
k_{,x_i} = \frac{-\beta}{2} k r^2_{,x_i} = -\beta k X_i \tag{16}
$$

$$
k_{,y_h} = \beta k X_h \tag{17}
$$

$$k_{,x_i x_j} = \frac{-\beta}{2}(k_{,x_j} r^2_{,x_i} + k r^2_{,x_i x_j}) \tag{18}$$
$$k_{,x_i x_j y_h} = \frac{-\beta}{2}(k_{,}x_j y_h r^2_{,x_i} + k_{,x_j} r^2_{,x_i y_h} + k_{,y_h} r^2_{,x_i x_j}) \tag{19}$$
$$k_{,x_i x_j y_h y_k} = \frac{-\beta}{2}(k_{,x_j y_h y_k} r^2_{,x_i} + k_{,x_j y_h} r^2_{,x_i y_k} + k_{,x_j y_k} r^2_{,x_i y_h} + k_{,y_h y_k r^2_{,x_i x_j}}) \tag{20}$$

second derivatives

just work on terms that are parts of the 4th kernel derivative

$$\begin{aligned}
k_{,x_i x_j} &= \frac{\partial}{\partial x_j}(-\beta k X_i) \\
&= -\beta(k_{,x_j X_i} + k X_{i,x_j}) \\
&= -\beta(-\beta k X_j X_i + k \delta_{ij} D_{ij}) \\
&= (\beta^2 X_j X_i - \beta \delta_{ij} D_{ij}) k
\end{aligned}$$

$$\begin{aligned}
k_{,x_i y_h} &= \frac{\partial}{\partial y_h}(-\beta k X_i) \\
&= -\beta(k_{,y_h X_i} + k X_{i,y_h}) \\
&= -\beta(\beta k X_h X_i - k \delta_{ih} D_{ih}) \\
&= -(\beta^2 X_h X_i - \beta \delta_{ih} D_{ih}) k
\end{aligned}$$

$$\begin{aligned}
k_{,y_h y_k} &= \frac{\partial}{\partial y_h}(\beta k X_h) \\
&= \beta(k_{,y_k X_h} + k X_{h,y_k}) \\
&= \beta(\beta k X_k X_h - k \delta_{hk} D_{hk}) \\
&= (\beta^2 X_h X_k - \beta \delta_{hk} D_{hk}) k
\end{aligned}$$

$$k_{,x_i x_j} = (\beta^2 X_j X_i - \beta \delta_{ij} D_{ij}) k \tag{21}$$
$$k_{,x_i y_h} = -(\beta^2 X_h X_i - \beta \delta_{ih} D_{ih}) k \tag{22}$$
$$k_{,y_h y_k} = (\beta^2 X_h X_k - \beta \delta_{hk} D_{hk}) k \tag{23}$$

Term 1 of the 4th derivative in eqn (20)

$$\begin{aligned}
k_{,x_j y_h y_k} &= \frac{\partial}{\partial y_k} k_{,x_j y_h} \\
&= -\frac{\partial}{\partial y_k}(\beta^2 X_h X_j - \beta \delta_{jh} D_{jh}) k \\
&= (\beta^2 D_{hk} \delta_{hk} X_j + \beta^2 X_h D_{jk} \delta_{jk}) k - (\beta^2 X_h X_j - \beta D_{jh} \delta_{jh}) \beta X_k k
\end{aligned}$$

$$
\begin{aligned}
&k_{,x_j y_h y_k} r^2_{,x_i} \\
&= 2[\beta^2 X_j D_{hk}\delta_{hk} + \beta^2 X_h D_{jk}\delta_{jk} + \beta^2 X_k D_{jh}\delta_{jh} - \beta^3 X_h X_j X_k] X_i k \\
&= \boxed{2\beta^2[X_j X_i D_{hk}\delta_{hk} + X_h X_i D_{jk}\delta_{jk} + X_k X_i D_{jh}\delta_{jh}]k - 2\beta^3 X_h X_j X_k X_i k}
\end{aligned}
$$

Term 2 of the 4th derivative

$$
\begin{aligned}
k_{,x_j y_h} r^2_{,x_i y_k} &= -(\beta^2 X_h X_j - \beta D_{jh}\delta_{jh})(-2D_{ik}\delta_{ik} k) \\
&= \boxed{(-2\beta^2 X_h X_j D_{ik}\delta_{ik} + 2\beta D_{jh} D_{ik}\delta_{jh}\delta_{ik})k}
\end{aligned}
$$

Term 3 of the 4th derivative

This is completely analogous to term 2 except the subscripts are slightly different

$$
k_{,x_j y_k} r^2_{,x_i y_h} = \boxed{(-2\beta^2 X_k X_j D_{ih}\delta_{ih} + 2\beta D_{jk} D_{ih}\delta_{jk}\delta_{ih})k}
$$

Term 4 of the 4th derivative

$$
\begin{aligned}
k_{,y_h y_k} r^2_{,x_i x_j} &= (\beta^2 X_k X_h - \beta D_{hk}\delta_{hk}) k 2 D_{ij}\delta_{ij} \\
&= \boxed{(2\beta^2 X_k X_h D_{ij}\delta_{ij} - 2\beta D_{ij} D_{hk}\delta_{hk}\delta_{ij})k}
\end{aligned}
$$

Collect terms of $\Sigma_{,x_i x_j y_h y_k}$ by plugging them in eqn 20

All the relevant terms are boxed above, note that we have to figure out the signs for the permutation terms

$$
\Sigma_{,x_i x_j y_h y_k} = (\beta^4 X_h X_j X_k X_i - \beta^3 (X_j X_i D_{hk}\delta_{hk} + 5\text{perm.}) - \beta^2 (D_{jh} D_{ik}\delta_{jh}\delta_{ik} + 2\text{perm.}))\Sigma \tag{24}
$$

Test 1:

Let's check that our general expression of the 4th derivative of Σ is correct by working out an example

notes

- γ_2, unlike κ and γ_1 does not have any pair of repeated indices, e.g. 1122, nor 2211 nor 1111 etc., so for small angular separation, only κ and γ_1 has increased covariances on the diagonal compared to ψ_s

Thoughts on implementation

- The metric object should incorporate the δ_{ij} condition for diagonal D, which will kill a lot of terms

Comparison between parametrization of George and our parametrization