

- **TITLE**

Parametrized Covariance Modeling for non-Homogeneous (and non-Stationary) Spatio-Temporal Random Process on the Sphere

- **ABSTRACT**

Identifying an appropriate covariance function is one of the primary interests in spatial or spatio-temporal data analysis in that it allows researchers to analyze the dependence structure and predict unobserved values of the process. For this purpose, homogeneity is a widely used assumption in spatial or spatio-temporal statistics, and many parameterized covariance models have been developed under this assumption. However, this is a strong and unrealistic condition in many cases. In addition, although different statistical approaches should be applied to build a proper covariance model on the sphere considering its unique characteristics, relevant studies are relatively less common. In this research, we introduce a novel parametrized model of the covariance function for non-homogeneous (and non-stationary) spatio-temporal random process on the sphere. To alleviate the homogeneity assumption and consider its spherical domain, this research applies the theories of intrinsic random function (IRF) while considering the significant influence of time components in the model as well. We also provide a methodology to estimate the parameters of intrinsic covariance function (ICF) that has a key role for prediction through kriging. Finally, the simulation study demonstrates validity of the suggested covariance model with its advantage of interpretability.

- **Keywords** Non-homogeneity, Non-stationarity, Spatio-temporal statistics, Covariance function, Sphere, Intrinsic Random Functions

- Why do we need the finite second moment for $X(P, t)$? Is it to guarantee the existence of covariance?

- How can we justify our mom estimator without ergodicity?

- **Ergodicity**

I think this definition is more straight forward than that of Cressie (P55)

If $\omega[t]$ eventually visits all of Ω regardless of $\omega[0]$, then Birkoff's equality (1931) holds :

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(\omega[t]) dt = \int_{\Omega} f(\omega) P(\omega) d\omega$$

The left one is average of a long trajectory (called time average) and the right one is average over all possible states (also called ensemble average)

- Think about AR(1), $X_n(t+1) = \phi X_n(t) + \epsilon(t)$ with $|\phi| < 1$

If we use an autoregressive model, we can use the ergodicity assumption since when $|\phi| < 1$, it forgets initial condition over time.

If initial conditions are very influential, that random process is not ergodic.

- Ergodic time series has to be "strongly stationary" (a second-order stationarity is not sufficient). However, statisticians are often using only part of the ergodicity assumption to guarantee the convergences of the sample mean and covariance to their population.(Cressie, p57).

This notion formulated by Gardiner (1983) is ergodicity in mean and ergodicity in covariance. These are specified by L_2 convergence of the sample quantities. (i.e., $E(X_n - X)^2 \rightarrow 0$ as $n \rightarrow \infty$). He also gives sufficient conditions for convergence that depends on fourth-order moments of the process.????

- For Gaussian random process, second-order stationarity and strong stationarity coincide because the distribution is specified by its mean and covariance. A sufficient condition for ergodicity is $C(h) \rightarrow 0$ as $\|h\| \rightarrow \infty$ (Adler, 1981).

- **Nugget Effect**

$$\gamma(-h) = \gamma(h) \text{ and } \gamma(0) = 0.$$

If $\gamma(h) \rightarrow C_0 > 0$ as $h \rightarrow 0$, then c_0 has been called the **nugget effect** (Matheron, 1962).

- It is believed that microscale variation (small nuggets) is causing a discontinuity at the origin.

$$C_0 = C_{MS} + C_{ME}$$

C_{MS} is a measurement error variance. C_{ME} is a white noise.

- The behavior of the variogram near the origin is very informative about the continuity properties of the random process $Z(\cdot)$. According to Matheron(1971b, p58):

1. $2\gamma(\cdot)$ is continuous at the origin. Then $Z(\cdot)$ is L_2 -continuous. [Clearly, $E(Z(s+h) - Z(s))^2 \rightarrow 0$ iff $2\gamma(h) \rightarrow 0$, as $\|h\| \rightarrow 0$.]

2. $2\gamma(\cdot)$ does not approach 0 as h approaches the origin. Then $Z(\cdot)$ is not even L_2 -continuous and is highly irregular. This discontinuity of γ at the origin is the **nugget effect** discussed previously.

3. $2\gamma(\cdot)$ is a positive constant (except at the origin where it is zero). Then $Z(s_1)$ and $Z(s_2)$ are uncorrelated for any $s_1 \neq s_2$, regardless of how close they are; $Z(\cdot)$ is often called white noise.

- The classical variogram estimator is unbiased for $2\gamma(\cdot)$ when $Z(\cdot)$ is intrinsically stationary. However, when $Z(\cdot)$ is second-order stationary, \hat{C} has $O(1/n)$ bias: $E(\hat{C}) = C(h) + O(1/n)$.

$$\hat{C} = \frac{1}{|N(h)|} \sum_{N(h)} (Z(s_i) - \bar{Z})(Z(s_j) - \bar{Z}) \quad \text{where} \quad \bar{Z} = \sum_{i=1}^n Z(s_i)/n$$

$$2\hat{\gamma}(h) \equiv \frac{1}{|N(h)|} \sum_{N(h)} (Z(s_i) - Z(s_j))^2 \quad \text{where} \quad N(h) = \{(s_i, s_j) : s_i - s_j = h; i, j = 1, \dots, n\}$$

- Variogram is unbiased but covariogram(covariance) is biased??? Probably, \bar{Z} is unbiased (if we can assume ergodicity) but \bar{Z}^2 might be biased.

- **Classical Variogram Estimator** $2\hat{\gamma}$ (Cressie, p96)

Assuming a Gaussian model,

$$\{Z(s+h) - Z(s)\}^2 \sim 2\gamma(h) \cdot \chi_1^2$$

$$E\left(\{Z(s+h) - Z(s)\}^2\right) = 2\gamma(h)$$

$$\text{var}\left(\{Z(s+h) - Z(s)\}^2\right) = 2(2\gamma(h))^2$$

$$\begin{aligned} & \text{corr}\left(\{Z(s_1+h_1) - Z(s_1)\}^2, \{Z(s_2+h_2) - Z(s_2)\}^2\right) \\ &= \frac{\left\{ \gamma(s_1 - s_2 + h_1) + \gamma(s_1 - s_2 - h_2) - \gamma(s_1 - s_2 + h_1 - h_2) - \gamma(s_1 - s_2) \right\}^2}{2\gamma(h_1) \cdot 2\gamma(h_2)} \end{aligned}$$

Thus, we can compute $\text{var}(2\hat{\gamma}(h(j)))$ and $\text{cov}(2\hat{\gamma}(h(i)), 2\hat{\gamma}(h(j)))$, which allows $V(\theta)$. Then, the weighted-least-squares criterion becomes:

$$(2\hat{\gamma} - 2\gamma(\theta))' V(\theta)^{-1} (2\hat{\gamma} - 2\gamma(\theta))$$

which can be a complicated function of θ to minimize.

Cressie(1985a) suggested that

$$\sum_{j=1}^K |N(h(j))| \left\{ \frac{\gamma(\hat{h}(j))}{\gamma(h(j); \theta)} - 1 \right\}^2$$

is a good approximation o w.l.s.

”This criterion is sensible from the viewpoint that the more pairs of observations $|N(h(j))|$ there are the more weight the residual at lag $h(j)$ receives in the overall fit. Also, the smaller the value of the theoretical variogram, the more weight the residual receives at the lag (i.e., lags closest to $h = 0$ typically get more weight, which is an attractive property because it is important to obtain a good fit of the variogram near the origin; see Stein, 1988). This criterion could be seen as a pragmatic compromise between efficiency (generalized least squares) and simplicity (ordinary least squares).”

- variogram is smaller at $h = 0$ unlike ICF???

Then, our model and criterion have wrong(or opposite) weights?

- Bessel Function. Matern covariance function.
- relationship b/w variogram and ICF???
- spectral representation of a variogram?
- Is our covariance function is isotropic?

Jan 22, 2023

- [check the DA paper and add more detail here.](#)

Let $X(P, t)$ be a spatio-temporal random process on the sphere and be continuous in quadratic mean. Then, we can expand such a random process through its spectral representation, which is convergent in quadratic mean. (Yaglom, 1961, Jones, 1963, Roy, 1969)

$$X(P, t) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t) Y_{\ell}^m(P)$$

$$Z_{\ell,m}(t) = \int_{\mathbb{S}^2} X(P, t) Y_{\ell}^m(P) dP$$

The $Y_{\ell}^m(\cdot)$ s are spherical harmonics, which are orthonormal basis functions of the sphere and do not depend on time $t \in \mathbb{R}$. On the other hand, each coefficient $Z_{\ell,m}(t)$ is a function of a time term t and free from the location $P \in \mathbb{S}^2$ in that it is integrated in terms of P .

- [\(Add introduction of IRF, allowable measure and Nil space here\)](#) Now, let assume $X(P, t)$ is an $\text{IRF}(\kappa)/\text{I}(0)$. In other words, $X(P, t)$ is non-homogeneous for the spatial term but still stationary in terms of time component. Then, we can say that $X(P, t) = \sum_{\ell=0}^{\kappa-1} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t) Y_{\ell}^m(P) + X_{\kappa}(P, t)$ where $X_{\kappa}(P, t) = \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t) Y_{\ell}^m(P)$. By introducing the new basis of Nil space, $q_1(\cdot), q_2(\cdot), \dots, q_{\kappa^2}(\cdot) \in N$ such that $q_{\nu}(\tau_{\mu}) = I(\nu, \mu)$ for $1 \leq \nu, \mu \leq \kappa^2$, we can re-express that $X(P, t) = \sum_{\nu=0}^{\kappa^2} Z_{\nu}(t) q_{\nu}(P) + X_{\kappa}(P, t)$ (Huang et al. 2019). Huang(2018) showed that the low frequency truncated process $X_{\kappa}(P, t)$ is homogeneous if the original process $X(P, t)$ is an $\text{IRF}(\kappa)$ on the sphere. In addition, since $X_{\kappa}(P, t)$ is homogenous and stationary, according to Roy(1969), the stochastic process $\{Z_{\ell,m}(t) : t \in \mathbb{R}\}$ is stationary for all $\ell \geq \kappa$ and $-\ell < m < \ell$; also, they are uncorrelated for different ℓ and m , i.e., $\text{Cov}\left(Z_{\ell,m}(t), Z_{\ell',m'}(t')\right) = 0$ for $\ell \neq \ell'$ or $m \neq m'$ when $\ell, \ell' \geq \kappa$.

Considering these facts, we can get a covariance function of $X_{\kappa}(P, t)$ such that:

$$\begin{aligned}
& Cov\left(\sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t) Y_{\ell}^m(P), \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(s) Y_{\ell}^m(Q)\right) \\
&= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{\ell'=\kappa}^{\infty} \sum_{m'=-\ell'}^{\ell'} Cov\left(Z_{\ell,m}(t), Z_{\ell',m'}(s)\right) Y_{\ell}^m(P) Y_{\ell'}^{m'}(Q)
\end{aligned}$$

By shur's decomposition (Roy, 1969),

$$\begin{aligned}
&= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(P) Y_{\ell}^m(Q) \quad \text{where} \quad a_{\ell}(h) = Cov\left(Z_{\ell,m}(t), Z_{\ell,m}(s)\right), \quad h = t - s \\
&= \sum_{\ell=\kappa}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{PQ}) \quad \text{by addition theorem.} \\
&= \phi_{\kappa}(\overrightarrow{PQ}, h)
\end{aligned}$$

- $\phi_{\kappa}(\overrightarrow{PQ}, h)$ is called Intrinsic Covariance Function(ICF) with order κ . This is homogenous and stationary. (add importance of the ICF here)
- Now, we can derive the covariance function of $X(P, t)$.

•

$$\begin{aligned}
& Cov\left(X(P, t), X(Q, s)\right) \\
&= Cov\left(\sum_{\nu=0}^{\kappa^2} Z_{\nu}(t)q_{\nu}(P) + \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \sum_{\mu=0}^{\kappa^2} Z_{\mu}(s)q_{\mu}(Q) + \sum_{\ell'=\kappa}^{\infty} \sum_{m'=-\ell'}^{\ell'} Z_{\ell',m'}(s)Y_{\ell'}^{m'}(Q)\right) \\
&= Cov\left(\sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \sum_{\ell'=\kappa}^{\infty} \sum_{m'=-\ell'}^{\ell'} Z_{\ell',m'}(s)Y_{\ell'}^{m'}(Q)\right) \\
&+ Cov\left(\sum_{\ell=0}^{\kappa-1} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \sum_{\mu=0}^{\kappa^2} Z_{\mu}(s)q_{\mu}(Q)\right) \\
&+ Cov\left(\sum_{\nu=0}^{\kappa^2} Z_{\nu}(t)q_{\nu}(P), \sum_{\ell'=\kappa}^{\infty} \sum_{m'=-\ell'}^{\ell'} Z_{\ell',m'}(s)Y_{\ell'}^{m'}(Q)\right) \\
&+ Cov\left(\sum_{\nu=0}^{\kappa^2} Z_{\nu}(t)q_{\nu}(P), \sum_{\mu=0}^{\kappa^2} Z_{\mu}(s)q_{\mu}(Q)\right) \\
&= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{\ell'=\kappa}^{\infty} \sum_{m'=-\ell'}^{\ell'} Cov\left(Z_{\ell,m}(t), Z_{\ell',m'}(s)\right) Y_{\ell}^m(P) Y_{\ell'}^{m'}(Q) \\
&+ \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{\mu=0}^{\kappa^2} Cov\left(Z_{\ell,m}(t), Z_{\mu}(s)\right) Y_{\ell}^m(P) q_{\mu}(Q) \\
&+ \sum_{\nu=0}^{\kappa^2} \sum_{\ell'=\kappa}^{\infty} \sum_{m'=-\ell'}^{\ell'} Cov\left(Z_{\nu}(t), Z_{\ell',m'}(s)\right) q_{\nu}(P) Y_{\ell'}^{m'}(Q) \\
&+ \sum_{\nu=0}^{\kappa^2} \sum_{\mu=0}^{\kappa^2} Cov\left(Z_{\nu}(t), Z_{\mu}(s)\right) q_{\nu}(P) q_{\mu}(Q)
\end{aligned}$$

Since $X(P, t)$ is an $IRF(\kappa)/I(0)$, which is not homogenous, we cannot guarantee that the covariance functions related to the low frequency $Z_{\nu}(\cdot)$ and $Z_{\mu}(\cdot)$ are 0 for $1 \leq \nu, \mu \leq \kappa^2$. That is, it is plausible and even more reasonable to assume that $Cov(Z_{\ell,m}(t), Z_{\mu}(s))$, $Cov(Z_{\nu}(t), Z_{\ell',m'}(s))$, and $Cov(Z_{\nu}(t), Z_{\mu}(s))$ are not zero when $\ell, \ell' \geq \kappa$; $\ell \leq m \leq \ell$,

$\ell' \leq m' \leq \ell'$; and $t, s \in \mathbb{R}$. In other words, elements of Nil space $N = \{Z_\nu(t) : 1 \leq \nu \leq \kappa^2\}$ are correlated with the other coefficients in contrast to the previous case of $X_\kappa(P, t)$, which is homogenous. **CHECK WHETHER IT MAKES SENSE????** In fact, Huang(2016) showed that coefficients of the low frequency can be correlated with the other coefficients of higher frequencies by providing an example of the Brownian bridge, which is an IRF(1) on a circle. In this research, our goal is to introduce appropriate structures for these covariances functions of non-homogenous or non-stationary processes. In pursuit of this aim, we can assume that the covariance functions of the coefficients have structures such that :

$$\begin{aligned}
Cov\left(Z_{\ell,m}(t), Z_{\ell',m'}(s)\right) &= a_\ell(h) I\{(\ell, m), (\ell', m')\} \\
Cov\left(Z_\nu(t), Z_{\ell,m}(s)\right) &= Cov\left(Z_{\ell,m}(t), Z_\nu(s)\right) = -a_\ell(h) Y_\ell^m(\tau_\nu) \\
Cov\left(Z_\nu(t), Z_\mu(s)\right) &= I(\nu, \mu) + \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_\ell(h) Y_\ell^m(\tau_\nu) Y_\ell^m(\tau_\mu) \\
&= I(\nu, \mu) + \sum_{\ell=\kappa}^{\infty} \frac{2\ell+1}{4\pi} a_\ell(h) P_\ell(\cos \overrightarrow{\tau_\nu \tau_\mu})
\end{aligned}$$

Let $a_\ell(h) \geq 0$ be positive definite and be expressed with a form of the ℓ th power.

That is, $a_\ell(h) = g^\ell(h)$. Then,

$$= I(\nu, \mu) + \frac{1 - g(h)^2}{\left(1 - 2g(h) \cdot \cos(\overrightarrow{\tau_\nu \tau_\mu}) + g^2(h)\right)^{3/2}}$$

In fact, this covariance function is free from ℓ if $a_\ell(h)$ has a form of $g(h)^\ell$

(by addition theorem & the generating function of Legendre Polynomial),

and also free from m (by Shur's decomposition.) Does it make sense??????

How can we justify these structures and assumptions???? Restriction to guarantee positive definiteness. These structures allow positive definiteness to the covariance model. Can we use RK for fixed h to justify this assumption?

In the case of the reproducing kernel for fixed h , each covariance function of the coefficients are following (from note 101022, page 67):

$$1 \leq \mu, \nu \leq \kappa^2 \quad \ell_2, \ell_2' \geq \kappa$$

$$b_{\ell_2, m_2}(h) = a_{\ell_2}(h) \quad \text{where} \quad a_{\ell}(h) = p_1^{\ell} e^{-p_2 \ell |h|}, \quad 0 < p_1 < 1, \quad p_2 > 0, \quad \ell = 0, 1, 2, \dots$$

$$b_{\mu}^{\nu}(h) = I_{(\mu, \nu)}(\mu, \nu) + \sum_{\ell_2=\kappa}^{\infty} a_{\ell_2}(h) Y_{\ell_2}^{m_2}(\tau_{\mu}) Y_{\ell_2}^{m_2}(\tau_{\nu}) \quad \text{where} \quad \tau_{\mu}, \tau_{\nu} \in \mathbb{S}^2$$

$$b_{\mu}^{\ell_2, m_2}(h) = b_{\ell_2, m_2}^{\mu}(h) = -a_{\ell_2}(h) Y_{\ell_2}^{m_2}(\tau_{\mu})$$

Does it mean if we use the same structures as the RK, we can guarantee that such forms of covariances functions exists?

Then, by Shur's decomposition (Roy 1969), (Do we need more detailed explanation about Shur's decomposition here?)

$$\begin{aligned}
Cov\left(X(P, t), X(Q, s)\right) &= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(P) Y_{\ell}^m(Q) \\
&+ \sum_{\nu=1}^{\kappa^2} \sum_{\mu=1}^{\kappa^2} q_{\nu}(P) q_{\mu}(Q) \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_{\nu}) Y_{\ell}^m(\tau_{\mu}) + \sum_{\nu=1}^{\kappa^2} q_{\nu}(P) q_{\nu}(Q) \\
&- \sum_{\nu=1}^{\kappa^2} q_{\nu}(P) \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_{\nu}) Y_{\ell}^m(Q) \\
&- \sum_{\nu=1}^{\kappa^2} q_{\nu}(Q) \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(P) Y_{\ell}^m(\tau_{\nu})
\end{aligned}$$

By addition theorem,

$$\begin{aligned}
&= \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{PQ}) - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{PQ}) \right\} \\
&+ \sum_{\nu=1}^{\kappa^2} \sum_{\mu=1}^{\kappa^2} q_{\nu}(P) q_{\mu}(Q) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{\tau_{\nu}\tau_{\mu}}) - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{\tau_{\nu}\tau_{\mu}}) \right\} + \sum_{\nu=1}^{\kappa^2} q_{\nu}(P) q_{\nu}(Q) \\
&- \sum_{\nu=1}^{\kappa^2} q_{\nu}(P) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{Q\tau_{\nu}}) - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{Q\tau_{\nu}}) \right\} \\
&- \sum_{\nu=1}^{\kappa^2} q_{\nu}(Q) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{P\tau_{\nu}}) - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{P\tau_{\nu}}) \right\} \\
&= \phi_{\kappa}(\overrightarrow{PQ}, h) + \sum_{\nu=1}^{\kappa^2} \sum_{\mu=1}^{\kappa^2} \phi_{\kappa}(\overrightarrow{\tau_{\nu}\tau_{\mu}}, h) q_{\nu}(P) q_{\mu}(Q) + \sum_{\nu=1}^{\kappa^2} q_{\nu}(P) q_{\nu}(Q) \\
&- \sum_{\nu=1}^{\kappa^2} \phi_{\kappa}(\overrightarrow{Q\tau_{\nu}}, h) q_{\nu}(P) - \sum_{\nu=1}^{\kappa^2} \phi_{\kappa}(\overrightarrow{P\tau_{\nu}}, h) q_{\nu}(Q)
\end{aligned}$$

- (Add why we need the scale parameters and benefit of them here) The scale parameters $\gamma, \gamma_\nu, \gamma_\mu > 0$ can be introduced to achieve more flexibility.

Then, since we can express $a_\ell(h) = g^\ell(h)$, $\ell = 0, 1, 2, \dots$

$$\begin{aligned}
&\Rightarrow \gamma^2 \left\{ \frac{1 - g(h)^2}{(1 - 2 \cos(\overrightarrow{PQ})g(h) + g(h)^2)^{3/2}} - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_\ell(h) P_\ell(\cos \overrightarrow{PQ}) \right\} \\
&+ \sum_{\nu=1}^{\kappa^2} \sum_{\mu=1}^{\kappa^2} (\gamma_\nu \gamma_\mu) q_\nu(P) q_\mu(Q) \left\{ \frac{1 - g(h)^2}{(1 - 2(\cos \overrightarrow{\tau_\nu \tau_\mu})g(h) + g(h)^2)^{3/2}} - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_\ell(h) P_\ell(\cos \overrightarrow{\tau_\nu \tau_\mu}) \right\} \\
&+ \sum_{\nu=1}^{\kappa^2} \gamma_\nu^2 q_\nu(P) q_\nu(Q) \\
&- \sum_{\nu=1}^{\kappa^2} (\gamma \gamma_\nu) q_\nu(P) \left\{ \frac{1 - g(h)^2}{(1 - 2 \cos(\overrightarrow{\tau_\nu Q})g(h) + g(h)^2)^{3/2}} - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_\ell(h) P_\ell(\cos \overrightarrow{Q\tau_\nu}) \right\} \\
&- \sum_{\nu=1}^{\kappa^2} (\gamma \gamma_\nu) q_\nu(Q) \left\{ \frac{(1 - g(h)^2)}{(1 - 2 \cos(\overrightarrow{P\tau_\nu})g(h) + g(h)^2)^{3/2}} - \sum_{\ell=0}^{\kappa-1} \frac{2\ell+1}{4\pi} a_\ell(h) P_\ell(\cos \overrightarrow{P\tau_\nu}) \right\}
\end{aligned}$$

where $\gamma, \gamma_\nu > 0$, and $\tau_\nu, \tau_\mu \in \mathbb{S}^2$

- (If $\kappa = 1$, doesn't it look weird to have the constant terms, $-\frac{1}{4\pi} - \frac{1}{16\pi^2} - \frac{1}{4\pi^{\frac{3}{2}}}$ in the covariance function? How can we explain or justify this?)

1. This is cancelled if we use the RK form!!!! (This may be the reason why we need the minus signs for the 3rd and 4th components).

2. This one happens since we have the form $a_\ell(h) = g(h)^\ell$. So, this form probably has a more important role than we thought before.

3. Then, we should use the same scale parameters. Check the scale parameter γ_ν .

4. Also, it still has a constant term, "+1" from $\sum_{\nu=1}^{\kappa^2} q_\nu(P) q_\nu(Q)$. Is it to make

the kernel strictly positive definite???

- Now, we want to verify the positive definiteness of the covariance function. To prove the positive definiteness for this, we need to show:

$$\sum_{i=1}^n \sum_{j=1}^n c_i \text{Cov}\left(X(y_i, t_i), X(y_j, t_j)\right) \bar{c}_j \geq 0 \quad \text{where} \quad y_i, y_j \in \mathbb{S}^2, \quad t_i, t_j \in \mathbb{R} \text{ or } \mathbb{Z} \quad c_i, c_j \in \mathbb{C}$$

Proof.

$$\begin{aligned}
& \sum_{i=1}^n \sum_{j=1}^n c_i \text{Cov} \left(X(y_i, t_i), X(y_j, t_j) \right) \bar{c}_j \\
&= \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \left\{ \gamma^2 \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell_2}^{\ell_2} a_{\ell}(h) Y_{\ell}^m(y_i) Y_{\ell}^m(y_j) + \sum_{\nu=1}^{\kappa^2} \sum_{\mu=1}^{\kappa^2} (\gamma_{\nu} \gamma_{\mu}) q_{\nu}(y_i) q_{\mu}(y_j) \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_{\nu}) Y_{\ell}^m(\tau_{\mu}) \right. \\
&+ \sum_{\nu=1}^{\kappa^2} \gamma_{\nu}^2 q_{\nu}(y_i) q_{\nu}(y_j) - \sum_{\nu=1}^{\kappa^2} (\gamma \gamma_{\nu}) q_{\nu}(y_i) \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_{\nu}) Y_{\ell}^m(y_j) - \sum_{\nu=1}^{\kappa^2} (\gamma \gamma_{\nu}) q_{\nu}(y_j) \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(y_i) Y_{\ell}^m(\tau_{\nu}) \left. \right\} \\
&= \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) \left\{ \gamma Y_{\ell}^m(y_i) - \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_i) Y_{\ell}^m(\tau_{\nu}) \right\} \left\{ \gamma Y_{\ell}^m(y_j) - \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_j) Y_{\ell}^m(\tau_{\nu}) \right\} \\
&+ \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \sum_{\nu=1}^{\kappa^2} \gamma_{\nu}^2 q_{\nu}(y_i) q_{\nu}(y_j)
\end{aligned}$$

By Bochner theorem,

$$\begin{aligned}
&= \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} \int_{\mathbb{R}} e^{i\omega(t_i - t_j)} F_{\ell, m}(d\omega) \left\{ \gamma Y_{\ell}^m(y_i) + \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_i) Y_{\ell}^m(\tau_{\nu}) \right\} \left\{ \gamma Y_{\ell}^m(y_j) + \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_j) Y_{\ell}^m(\tau_{\nu}) \right\} \\
&+ \sum_{i=1}^n \sum_{j=1}^n c_i \bar{c}_j \sum_{\nu=1}^{\kappa^2} \gamma_{\nu}^2 q_{\nu}(y_i) q_{\nu}(y_j)
\end{aligned}$$

where $h_{ij} = t_i - t_j$, and $F_{\ell, m}(d\omega)$ is a non-negative measure.

$$\begin{aligned}
&= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} \int_{\mathbb{R}} F_{\ell, m}(d\omega) \sum_{i=1}^n c_i e^{i\omega t_i} \left\{ \gamma Y_{\ell}^m(y_i) + \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_i) Y_{\ell}^m(\tau_{\nu}) \right\} \sum_{j=1}^n \bar{c}_j e^{-i\omega t_j} \left\{ \gamma Y_{\ell}^m(y_j) + \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_j) Y_{\ell}^m(\tau_{\nu}) \right\} \\
&+ \sum_{\nu=1}^{\kappa^2} \left\{ \sum_{i=1}^n c_i \gamma_{\nu} q_{\nu}(y_i) \right\} \left\{ \sum_{j=1}^n \bar{c}_j \gamma_{\nu} q_{\nu}(y_j) \right\} \\
&= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} \int_{\mathbb{R}} F_{\ell, m}(d\omega) \left| \sum_{i=1}^n c_i e^{i\omega t_i} \left\{ \gamma Y_{\ell}^m(y_i) + \sum_{\nu=1}^{\kappa^2} \gamma_{\nu} q_{\nu}(y_i) Y_{\ell}^m(\tau_{\nu}) \right\} \right|^2 + \sum_{\nu=1}^{\kappa^2} \left\{ \sum_{i=1}^n c_i \gamma_{\nu} q_{\nu}(y_i) \right\}^2 \geq 0
\end{aligned}$$

Therefore, $\text{Cov} \left(X(y_i, t_i), X(y_j, t_j) \right)$ is positive-semi definite.

□

- Now, for the sake of simplicity, let consider IRF(1)/I(0), i.e, $\kappa = 1$. Then, its covariance function is:

$$\begin{aligned}
Cov\left(X(P, t), X(Q, s)\right) &= Cov\left(Z_{0,0}(t)Y_0^0(P) + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \quad Z_{0,0}(s)Y_0^0(Q) + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(s)Y_{\ell}^m(Q)\right) \\
&= Cov\left(Z_0(t)q_1(t)(P) + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \quad Z_0(s)q_1(Q) + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(s)Y_{\ell}^m(Q)\right) \\
&= Cov\left(\sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \quad \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(s)Y_{\ell}^m(Q)\right) \\
&\quad + Cov\left(Z_0(t)q_1(P), \quad Z_0(s)q_1(Q)\right) \\
&\quad + Cov\left(Z_0(t)q_1(P), \quad \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(s)Y_{\ell}^m(Q)\right) \\
&\quad + Cov\left(\sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Z_{\ell,m}(t)Y_{\ell}^m(P), \quad Z_0(s)q_1(Q)\right) \\
&= \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{\ell'=1}^{\infty} \sum_{m'=-\ell'}^{\ell'} Cov\left(Z_{\ell,m}(t), Z_{\ell',m'}(s)\right) Y_{\ell}^m(P) Y_{\ell'}^{m'}(Q) \\
&\quad + Cov\left(Z_0(t), Z_0(s)\right) q_1(P) q_1(Q) \\
&\quad + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Cov\left(Z_0(t), Z_{\ell,m}(s)\right) q_1(P) Y_{\ell}^m(Q) \\
&\quad + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} Cov\left(Z_{\ell,m}(t), Z_0(s)\right) Y_{\ell}^m(P) q_1(Q)
\end{aligned}$$

Let

$$\begin{aligned}
Cov\left(Z_{\ell,m}(t), Z_{\ell',m'}(s)\right) &= a_{\ell}(h) I\{(\ell, m), (\ell', m')\} \\
Cov\left(Z_0(t), Z_{\ell,m}(s)\right) &= Cov\left(Z_{\ell,m}(t), Z_0(s)\right) = -a_{\ell}(h) Y_{\ell}^m(\tau_1) \quad \text{where} \quad \tau_1 \in \mathbb{S}^2 \\
Cov\left(Z_0(t), Z_0(s)\right) &= \sum_{\ell=\kappa}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_1) Y_{\ell}^m(\tau_1) + 1 \\
&= 1 + \sum_{\ell=\kappa}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos 0) = 1 + \sum_{\ell=\kappa}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h)
\end{aligned}$$

Let $a_{\ell}(h) \geq 0$ be positive definite and be expressed with a form of the ℓ th power.

That is, $a_{\ell}(h) = g^{\ell}(h)$. Then,

$$= I(\nu, \mu) + \frac{1 - g(h)^2}{\left(1 - 2g(h) + g^2(h)\right)^{3/2}}$$

In fact, this covariance function is free from ℓ if $a_{\ell}(h)$ has a form of $g(h)^{\ell}$

(by addition theorem & the generating function of Legendre Polynomial),

and also free from m (by Shur's decomposition.) Does it make sense?????

Then, by Shur's decomposition (Roy 1969),

$$\begin{aligned}
&Cov\left(X(P, t), X(Q, s)\right) \\
&= \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(P) Y_{\ell}^m(Q) + \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_1) Y_{\ell}^m(\tau_1) q_1(P) q_1(Q) + q_1(\tau_1) q_1(\tau_1) \\
&\quad - q_1(P) \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(\tau_1) Y_{\ell}^m(Q) - q_1(Q) \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell}(h) Y_{\ell}^m(P) Y_{\ell}^m(\tau_1)
\end{aligned}$$

By addition theorem,

$$\begin{aligned}
&= \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{PQ}) - \frac{1}{4\pi} \right\} + q_1(P)q_1(Q) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) - \frac{1}{4\pi} \right\} + q_1(\tau_1)q_1(\tau_1) \\
&- q_1(P) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{Q\tau_1}) - \frac{1}{4\pi} \right\} - q_1(Q) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{P\tau_1}) - \frac{1}{4\pi} \right\} \\
&= \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{PQ}) - \frac{1}{4\pi} \right\} + \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) - \frac{1}{4\pi} \right\} + 1 \\
&- \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{Q\tau_1}) - \frac{1}{4\pi} \right\} - \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{P\tau_1}) - \frac{1}{4\pi} \right\} \\
&= \phi_1(\overrightarrow{PQ}, h) + \phi_1(0, h) + 1 - \phi_1(\overrightarrow{Q\tau_1}, h) - \phi_1(\overrightarrow{P\tau_1}, h)
\end{aligned}$$

After adding the scale parameters,

$$\begin{aligned}
&= \gamma_0^2 \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{PQ}) - \frac{1}{4\pi} \right\} + \gamma_1^2 \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) - \frac{1}{4\pi} \right\} + \gamma_1^2 \\
&- (\gamma_0 \cdot \gamma_1) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{Q\tau_1}) - \frac{1}{4\pi} \right\} - (\gamma_0 \cdot \gamma_1) \left\{ \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} a_{\ell}(h) P_{\ell}(\cos \overrightarrow{P\tau_1}) - \frac{1}{4\pi} \right\}
\end{aligned}$$

Since we can express $a_{\ell}(h) = g^{\ell}(h)$, $\ell = 0, 1, 2, \dots$,

$$\begin{aligned}
&\Rightarrow \gamma_0^2 \left\{ \frac{1 - g(h)^2}{(1 - 2 \cos(\overrightarrow{PQ})g(h) + g(h)^2)^{3/2}} - \frac{1}{4\pi} \right\} + \gamma_1^2 \left\{ \frac{1 - g(h)^2}{(1 - 2g(h) + g(h)^2)^{3/2}} - \frac{1}{4\pi} \right\} + \gamma_1^2 \\
&- (\gamma_0 \cdot \gamma_1) \left\{ \frac{1 - g(h)^2}{1 - 2 \cos(\overrightarrow{\tau_1 Q})g(h) + g(h)^2)^{3/2}} - \frac{1}{4\pi} \right\} - (\gamma_0 \cdot \gamma_1) \left\{ \frac{1 - g(h)^2}{(1 - 2 \cos(\overrightarrow{P\tau_1})g(h) + g(h)^2)^{3/2}} - \frac{1}{4\pi} \right\}
\end{aligned}$$

where $\gamma_0, \gamma_1 > 0$, $\tau_1 \in \mathbb{S}^2$

If $\gamma^{1/2} = \gamma_0 = \gamma_1 > 0$,

$$\begin{aligned}
&\Rightarrow \gamma \left\{ \frac{1 - g(h)^2}{(1 - 2 \cos(\overrightarrow{PQ})g(h) + g(h)^2)^{3/2}} - \frac{1 - g(h)^2}{1 - 2 \cos(\overrightarrow{\tau_1 Q})g(h) + g(h)^2)^{3/2}} - \frac{1 - g(h)^2}{(1 - 2 \cos(\overrightarrow{P\tau_1})g(h) + g(h)^2)^{3/2}} \right. \\
&\quad \left. + \frac{1 - g(h)^2}{(1 - 2g(h) + g(h)^2)^{3/2}} + 1 \right\}
\end{aligned}$$

- (If $\kappa = 1$, doesn't it look weird to have the constant terms, $-\frac{1}{4\pi} - \frac{1}{16\pi^2} - \frac{1}{4\pi^{\frac{3}{2}}}$ in the covariance function? How can we explain or justify this?)
 1. This is cancelled if we use the RK form!!!! (This may be the reason why we need the minus signs for the 3rd and 4th components).
 2. This one happens since we have the form $a_\ell(h) = g(h)^\ell$. So, this form probably has a more important role than we thought before.
 3. Then, we should use the same scale parameters. Check the scale parameter γ_ν .
 4. Also, it still has a constant term, "+1" from $\sum_{\nu=1}^{\kappa^2} q_\nu(P)q_\nu(Q)$. Is it to make the kernel strictly positive definite???

- **Simulation Study**

Simulation Result								
p_1	p_2	p_3	$Avg(\hat{p}_1)$	$Avg(\hat{p}_2)$	$Avg(\hat{p}_3)$	$sd(\hat{p}_1)$	$sd(\hat{p}_2)$	$sd(\hat{p}_3)$
0.95	0.005	1.00	0.9408	0.0085	1.2257	0.0161	0.0030	0.7512
0.90	0.001	1.00	0.9092	0.0015	0.7172	0.0238	0.0005	0.5741
0.90	0.40	1.00	0.8783	0.9655	1.4974	0.0081	0.6486	0.2626
0.90	0.90	1.00	0.8788	12.2837	1.4874	0.0080	8.8058	0.2652
p2 is bimodal								
0.85	0.01	1.00	0.8780	0.0126	0.6616	0.0339	0.0042	0.5787
0.80	0.10	0.50	0.8368	0.1280	0.4345	0.0416	0.0502	0.3505
0.75	0.15	10.0	0.8599	0.1412	14.6832	0.0703	0.1293	19.5012
0.70	0.05	1.00	0.8409	0.0577	1.1477	0.0910	0.0474	1.6579
0.70	0.20	1.00	0.8251	0.4521	67285.8498	0.0643	2.6349	249889.1306
0.70	0.50	1.00	0.8096	0.1862	47636.8859	0.0859	1.3436	250055.2191
0.65	0.20	10.0	0.7520	0.1262	439522.2402	0.1634	0.7395	1075563.7332
0.60	0.30	50.0	0.6598	0.1018	1046633.3154	0.2225	0.0457	2128092.1238
0.50	0.10	1.00	0.7449	0.1081	40059.3103	0.0459	0.5753	249243.1086
0.40	0.60	5.00	0.6714	0.1486	41854.6665	0.0605	0.0311	162109.1729
0.30	1.50	1.00	0.6549	0.1839	2227.7667	0.1176	0.6505	18415.3117
0.20	0.02	0.10	0.5218	0.0199	0.0167	0.0704	0.0155	0.0150
0.10	0.80	0.50	0.5776	0.1671	1.5545	0.0945	0.1403	1.4264

Table 1: Average values and standard deviations of 1,000 estimates with true parameter values for $IRF(2)/I(0)$. Each simulation includes 200 locations and 20 temporal points.

- How can we deal with the right skewed distribution. These might be failures of the optimization.