

Prueba 02 Análisis Estadístico de Datos Espaciales

José Flores

December 2024

1 Pregunta 1

1.1 Gráfico del Semivariograma Isotrópico

En la figura siguiente, se presenta el semivariograma isotrópico estimado a partir de los datos espaciales proporcionados. Este semivariograma se calculó utilizando el paquete `GeoModels` en R.

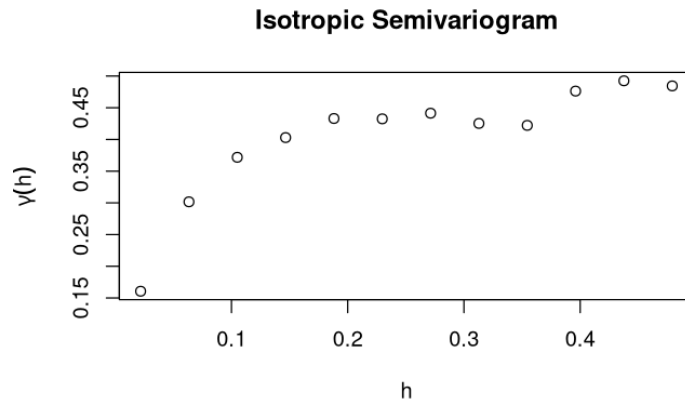


Figure 1: Semivariograma isotrópico estimado. El eje h representa las distancias, y el eje $\gamma(h)$ muestra la semivarianza.

Al asumir isotropía, se considera que la correlación entre puntos espaciales depende únicamente de la distancia h entre ellos y no de la dirección. Este supuesto simplifica la estimación del semivariograma, agrupando los pares de puntos por distancia en lugar de considerar orientaciones específicas.

En el semivariograma presentado:

- Se observa que la semivarianza $\gamma(h)$ crece rápidamente para valores pequeños de h , lo cual es típico de procesos espaciales con alta correlación a distancias cortas.
- Para valores mayores de h , $\gamma(h)$ parece estabilizarse, sugiriendo un alcance (*range*) más allá del cual los puntos no están correlacionados de manera "fuerte"
- No se detecta un *nugget* significativo (salto inicial en $h = 0$), lo que indica que la variabilidad a escalas muy pequeñas es baja o despreciable.

1.2 Distribución de los Datos

El histograma presentado en la Figura 2 muestra la distribución de los datos observados en las 500 coordenadas espaciales. Este servirá para justificar si es apropiado el uso del modelo Gaussiano según los datos que se tienen.

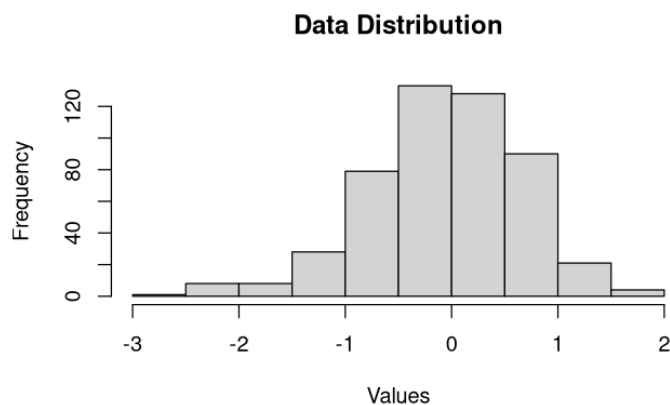


Figure 2: Distribución de los datos observados. El eje x representa los valores de los datos, y el eje y la frecuencia de ocurrencia.

Comentarios sobre el modelo Gaussiano:

- **Simetría:** El histograma es aproximadamente simétrico, lo que es indicativo de una distribución normal, consistente con la hipótesis del modelo Gaussiano.

- **Media y varianza:** Los datos están centrados alrededor de 0, lo cual respalda la suposición de que $E[Z(s)] = 0$ (la media real es de aproximadamente -0.057). Además, la dispersión aparentemente moderada de los valores es compatible con $Var(Z(s)) = 1$ (si bien la varianza real es aproximadamente de 0.51).
- **Ausencia de colas largas:** No se observan valores extremos o colas largas significativas, lo que respalda la normalidad de los datos.

Con base en este análisis visual, el modelo Gaussiano parece ser razonable para describir los datos observados.

1.3 Estimación de parámetros y su interpretación

Los parámetros estimados utilizando el método de máxima verosimilitud para los modelos de correlación son los siguientes:

- **Modelo 1 (Exponencial):**

Media (μ) : -0.0585,
 Escala (α) : 0.0804,
 Sill (σ^2) : 0.5182.

- **Modelo 2 (Matérn con $\nu = 3/2$):**

Media (μ) : -0.0581,
 Escala (α) : 0.0296,
 Sill (σ^2) : 0.5181.

1.3.1 Interpretación de los parámetros estimados

- **Media (μ):** En ambos modelos, la media estimada es cercana a 0, lo que es consistente con la media observada de los datos ($\mu = -0.057$). Esto indica que el ajuste de los modelos está alineado con las características de la distribución de los datos.
- **Sill (σ^2):** El sill representa la varianza total del proceso espacial. Los valores estimados ($\sigma^2 \approx 0.518$) coinciden con la varianza calculada directamente a partir de los datos (Varianza = 0.51). Esto confirma que ambos modelos capturan adecuadamente la variabilidad de los datos.

- **Escala (α):** El parámetro de escala controla la rapidez con la que disminuye la correlación entre los puntos espaciales al aumentar la distancia. Un valor más pequeño de α indica que la dependencia espacial se desvanece de manera más rápida:
 - En el modelo exponencial, $\alpha = 0.0804$, lo que sugiere que la correlación espacial se extiende moderadamente.
 - En el modelo Matérn con $\nu = 3/2$, $\alpha = 0.0296$, indicando una dependencia espacial más "local" en comparación con el modelo exponencial. Esto refleja que el modelo Matérn puede capturar estructuras espaciales de menor alcance (*range*) con mayor detalle. El modelo "recuerda" información de puntos más cercanos en comparación al exponencial.

La diferencia en los valores de α entre los modelos también está relacionada con la forma de la función de correlación de cada modelo: el modelo Matérn pareciera ser más flexible pudiendo ajustarse de mejor manera a estructuras espaciales complejas.

1.3.2 Conclusión

Ambos modelos presentan parámetros consistentes con las propiedades estadísticas de los datos. La elección entre ellos podría depender de criterios adicionales, como el criterio de información de Akaike (AIC) o la validación cruzada basada en el error cuadrático medio (RMSE) (¡spoilers!)

1.4 Selección del Mejor Modelo: Criterio de Información de Akaike (AIC)

El criterio de información de Akaike (AIC) es una medida que combina la "calidad de ajuste" de un modelo con su complejidad (definida por el número de parámetros del modelo). En general, un modelo con un AIC más bajo se considera mejor, ya que logra un equilibrio entre la calidad del ajuste y el número de parámetros.

Los valores de log-verosimilitud compuesta (`logCompLik`) obtenidos para los modelos son los siguientes:

- Modelo 1 (Exponencial): -542705.6 .
- Modelo 2 (Matérn con $\nu = 3/2$): -542992.5 .

El AIC se calcula como:

$$\text{AIC} = -2 \cdot \log\text{CompLik} + 2 \cdot k,$$

donde:

- $\log\text{CompLik}$ es el valor de log-verosimilitud compuesta del modelo.
- k es el número de parámetros estimados en el modelo.

Dado que ambos modelos tienen el mismo número de parámetros ($k = 3$: media, varianza y escala), la comparación del AIC depende únicamente de los valores de $-2 \cdot \log\text{CompLik}$. Por lo tanto, se selecciona el modelo con el $\log\text{CompLik}$ más alto (es decir, el menos negativo).

1.4.1 Interpretación de los Resultados

- Para el Modelo 1 (Exponencial): $-2 \cdot (-542705.6) = 1085411.2$.
- Para el Modelo 2 (Matérn con $\nu = 3/2$): $-2 \cdot (-542992.5) = 1085985.0$.

El Modelo 1 tiene un AIC más bajo, lo que indica que es el mejor modelo según el criterio de información de Akaike.

1.4.2 Conclusión

De acuerdo con el criterio de información de Akaike (AIC), el modelo exponencial es preferible al modelo Matérn ($\nu = 3/2$) para describir la estructura de dependencia espacial de los datos. Esto implica que la relación entre los datos espaciales se explica de manera más efectiva utilizando el modelo exponencial (el modelo exponencial podría generalizar mejor "datos no vistos")

1.5 Elección del Mejor Modelo: Criterio RMSE

El RMSE mide la diferencia entre los valores predichos y los valores observados. En particular, se realizó una validación cruzada con 100 iteraciones, utilizando el 10% de los datos como conjunto de validación en cada iteración.

Los valores promedio de RMSE obtenidos para ambos modelos fueron los siguientes:

- Modelo 1 (Exponencial): $\text{RMSE} = 0.4447$.
- Modelo 2 (Matérn, $\nu = 3/2$): $\text{RMSE} = 0.4743$.

Según la métrica de RMSE, el modelo con el menor RMSE tiene el mejor desempeño predictivo, por lo que se concluye que el **Modelo 1** (Exponencial) es el mejor modelo según el criterio RMSE, ya que tiene un RMSE menor que el Modelo 2 (Matérn). Esto indica que el Modelo 1 proporciona una mejor capacidad predictiva en comparación al de Matérn.

1.6 Gráficas de Kriging y MSE

A continuación se muestran los gráficos relacionados a la predicción (Kriging) y al MSE para cada punto:

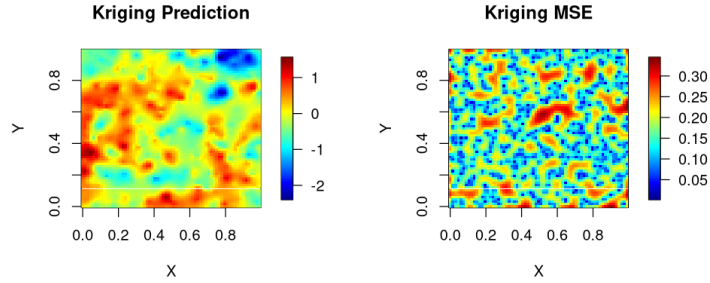


Figure 3: Mapa de calor de los puntos predichos. El eje x e y representan la coordenadas cartesianas correspondientes al eje X e Y, respectivamente. La escala del gráfico de la izquierda representa la predicción realizada y la de la derecha representa la magnitud del MSE.

1.6.1 Descripción Gráfica General

Pareciera que los puntos con el mayor MSE están asociados con los valores predichos más cercanos a -1 (en celeste). Esta relación puede sugerir que el modelo tiene una mayor incertidumbre (error) en las predicciones cuando los valores son más cercanos a -1. Esto podría estar relacionado con cómo el modelo maneja ciertas áreas del espacio donde los valores observados o predichos son "bajos", lo que genera una mayor dispersión o variabilidad en la predicción.

Por otro lado, el modelo presenta un gran número de puntos con valores cercanos a 1 y con MSE bajos (aproximadamente entre 0.05 y 0.10, en azul y celeste en el gráfico de MSE). Esto indica que en esas zonas el modelo es mucho más confiable y hace predicciones precisas. La baja varianza de MSE en esas áreas también sugiere que el modelo tiene una alta certeza en sus predicciones cuando los valores están cerca de 1.

1.6.2 Interpretación en los Gráficos

- **Gráfico de Predicciones (Kriging Prediction):** Las zonas continuas de color similar indican que el modelo está prediciendo valores similares en esas áreas. Esto sugiere que el modelo ha aprendido patrones espaciales donde la cercanía geográfica entre los puntos se traduce en similitudes en los valores predichos. Esto es una señal clara de que hay dependencia espacial en esas áreas.
- **Gráfico de MSE (Kriging MSE):** Similarmente, las zonas con colores continuos (y en particular las zonas de color azul o celeste, que indican bajos valores de MSE) también muestran áreas donde el modelo tiene alta certeza en sus predicciones. En estos casos, los puntos cercanos tienen un error bajo, lo que también sugiere que el modelo está capturando bien la dependencia espacial en esas zonas. Es decir, el modelo está ajustando adecuadamente los valores en función de la cercanía de los puntos.

1.6.3 Conclusión

Las zonas continuas de colores similares tanto en las predicciones como en el MSE indican que hay una fuerte dependencia espacial. Los puntos cercanos geográficamente tienen valores similares- (predicción) y errores similares (MSE). Esto refuerza la idea de que el modelo captura las correlaciones espaciales presentes en los datos, lo cual es consistente con la naturaleza de los campos aleatorios y los modelos de kriging.

2 Pregunta 2

2.1 Parámetros Iniciales y Restricciones

Antes de ajustar el modelo, se configuraron ciertos parámetros iniciales, que sirven como punto de partida para el proceso de optimización. Estos valores no son finales y se ajustarán durante el proceso de máxima verosimilitud. A continuación se detallan los parámetros utilizados:

- **Parámetros Iniciales:**

$$\begin{aligned}\hat{\mu} &= \text{media de la temperatura} \\ \hat{\mu}_1 &= 0 \text{ (Efecto inicial de la altitud)} \\ \hat{\alpha} &= 500 \text{ (Rango espacial inicial)} \\ \hat{\sigma}^2 &= \text{varianza de la temperatura}\end{aligned}$$

- **Parámetros Fijos para el Modelo Wendland:**

$$\text{power2} = 4 \text{ (Potencia del modelo de Wendland)}$$

- **Límites para los Parámetros:**

$$\text{Limite Inferior: } \mu = -\infty, \quad \mu_1 = -\infty, \quad \alpha = 0.01, \quad \sigma^2 = 0.01$$

$$\text{Limite Superior: } \mu = \infty, \quad \mu_1 = \infty, \quad \alpha = \infty, \quad \sigma^2 = \infty$$

Estos parámetros fueron elegidos de forma inicial para proporcionar un punto de partida en el proceso de optimización del modelo, y se ajustaron durante el cálculo de máxima verosimilitud.

2.2 Estimación de los Parámetros de Regresión Espacial

Se estimaron los parámetros β_0 , β_1 , σ^2 , y α utilizando el método de máxima verosimilitud compuesta basado en parejas condicionales, con una función de peso basada en vecindades de orden 3. Este método permite ajustar un modelo de regresión espacial teniendo en cuenta la dependencia espacial entre los datos. El modelo de regresión espacial utilizado es el siguiente:

$$Y(s) = \beta_0 + \beta_1 A(s) + \sigma Z(s)$$

Donde $Y(s)$ es la temperatura media en las coordenadas s , $A(s)$ es la altitud en esas ubicaciones, $Z(s)$ es un campo aleatorio Gaussiano con media 0 y varianza 1, y los parámetros a estimar son β_0 , β_1 , σ , y α , donde α es el parámetro de dependencia espacial. Los valores estimados (en R) para los parámetros son los siguientes:

$$\hat{\beta}_0 = 13.96, \quad \hat{\beta}_1 = -0.01, \quad \hat{\sigma} = 0.46, \quad \hat{\alpha} = 129.48$$

El parámetro $\hat{\beta}_1 = -0.01$ sugiere que la temperatura media disminuye ligeramente con el aumento de la altitud, aunque este efecto no es muy fuerte. La estimación de β_1 tiene un pequeño error estándar (ver inciso 2.2), lo que indica que la estimación es precisa.

2.3 Estimación de los Errores Estándar mediante el Método Bootstrap

El método bootstrap se utilizó para estimar los errores estándar de los parámetros ajustados. Este procedimiento se realizó a través de 500 simulaciones, y los errores estándar obtenidos son los siguientes:

- **Errores estándar:**

$$\begin{aligned}\hat{\beta}_0 &= 0.0685 \\ \hat{\beta}_1 &= 0.00008 \\ \hat{\sigma}^2 &= 1.11 \\ \hat{\alpha} &= 0.038\end{aligned}$$

Cabe destacar que los errores estándar son muy pequeños, lo que indica que la estimación es bastante precisa.

2.4 Comparación del Semivariograma Empírico de los Residuos con el Semivariograma Estimado

El semivariograma empírico de los residuos se calculó y se comparó con el semivariograma estimado del modelo ajustado. El semivariograma empírico muestra cómo varían los residuos en función de la distancia entre los puntos espaciales. Por otro lado, el semivariograma estimado es el que se obtiene a partir del modelo teórico de dependencia espacial ajustado. A continuación se muestran:

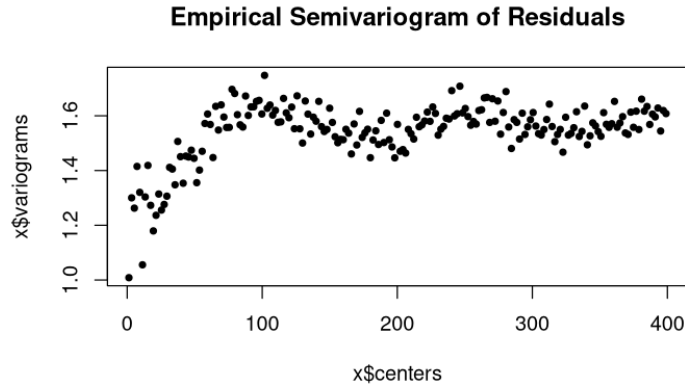


Figure 4: Semivariograma empírico de los residuos. Se eligió una distancia máxima de 400 unidades según la mitad de la distancia máxima entre los datos.

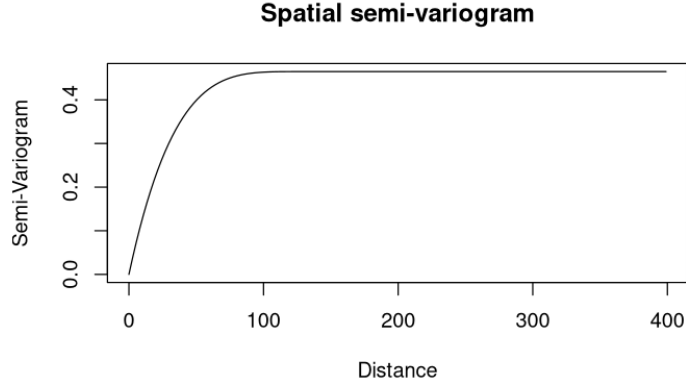


Figure 5: Comparación entre el semivariograma empírico y el semivariograma estimado y conclusiones

Se ve que ambos semivariogramas se empiezan a estabilizar desde el parámetro de escala estimado (alrededor de 129), el cual tiene sentido considerando el modelo de correlación y cómo la correlación desaparece para distancias mayores al parámetro de escala. Ambos semivariogramas se comportan de manera similar y se obtienen parámetros con bastante precisión, con lo que se concluye que el modelo es razonable.

3 Pregunta 3

3.1 Resultados Obtenidos

De acuerdo con el código y los resultados obtenidos, se pueden analizar los siguientes aspectos:

3.1.1 Parámetros Estimados

Los parámetros estimados para el modelo son los siguientes:

- **Media** $\mu = 0.0783556$
- **Nugget** $\text{nugget} = 2.146703 \times 10^{-11}$ (prácticamente 0)
- **Escala** $\alpha_s = 0.2040421$ (aproximadamente 0.2)
- **Sill** $\sigma^2 = 0.3278606$

El parámetro de **nugget** es muy pequeño, lo que sugiere que no hay variabilidad significativa a pequeñas distancias. Esto es apropiado si no existe dispersión excesiva en las distancias muy cortas.

3.1.2 Predicciones

Las predicciones obtenidas en las ubicaciones de interés son:

- **Primer punto de predicción:** -0.4219
- **Segundo punto de predicción:** 0.5391

Las varianzas de predicción para esos puntos son:

- **Primer punto de predicción:** 0.0364
- **Segundo punto de predicción:** 0.0422

Esto sugiere que el modelo tiene más certeza en la predicción del segundo punto, ya que la varianza es más baja en ese punto.

3.1.3 Gráfica de Ubicaciones de Predicción

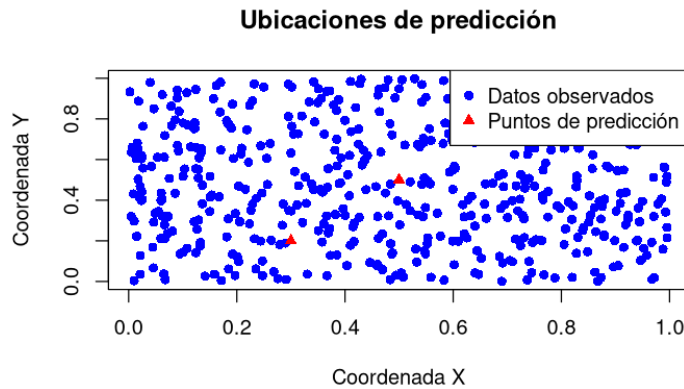


Figure 6: Ubicación de los puntos junto con sus predicciones. Ver leyendas.

La gráfica muestra que los puntos de predicción (rojos) están colocados sobre el conjunto de datos observados (azules), lo cual tiene sentido ya que las predicciones se realizaron en puntos específicos de la misma área geográfica.

3.2 Conclusiones

En resumen, los resultados sugieren lo siguiente:

- El modelo realiza predicciones razonables, y la mayor incertidumbre en el primer punto de predicción es esperada dada la mayor varianza de predicción.

4 Pregunta 4

4.1 Planteamiento

Sea $Z = \{Z(s), s \in A \subset \mathbb{R}^2\}$ un campo aleatorio Gaussiano con media constante $E(Z(s)) = \mu$, varianza unitaria $Var(Z(s)) = 1$ y función de correlación conocida $\rho(h)$. El objetivo es estimar μ mediante el método de máxima verosimilitud compuesta basado en parejas marginales, es decir, encontrar el estimador:

$$\hat{\mu} = \arg \max_{\mu \in \mathbb{R}} \left(\sum_{i=1}^n \sum_{j \neq i} l_{ij}(\mu) w_{ij} \right),$$

donde $l_{ij}(\mu)$ es la log-verosimilitud asociada al vector aleatorio bivariado Gaussiano $(Z(s_i), Z(s_j))^T$ y w_{ij} son pesos conocidos.

4.2 Desarrollo del Estimador

La log-verosimilitud marginal para un par $(Z(s_i), Z(s_j))$ está dada por:

$$l_{ij}(\mu) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(|\Sigma|) - \frac{1}{2} (Z_{ij} - \mu \mathbf{1})^T \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}),$$

donde:

- $Z_{ij} = \begin{pmatrix} Z(s_i) \\ Z(s_j) \end{pmatrix}$,
- $\mathbf{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,
- $\Sigma = \begin{pmatrix} 1 & \rho(h_{ij}) \\ \rho(h_{ij}) & 1 \end{pmatrix}$ es la matriz de covarianza,
- $\rho(h_{ij})$ es la correlación entre $Z(s_i)$ y $Z(s_j)$ dependiendo de la distancia $h_{ij} = \|s_i - s_j\|$.

Sustituyendo $l_{ij}(\mu)$ en la expresión del estimador, se obtiene:

$$\hat{\mu} = \arg \max_{\mu \in \mathbb{R}} \left(\sum_{i=1}^n \sum_{j \neq i} -\frac{1}{2} (Z_{ij} - \mu \mathbf{1})^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}) w_{ij} \right).$$

Ahora, se deriva la función objetivo respecto a μ . El término relevante es:

$$\sum_{i=1}^n \sum_{j \neq i} (Z_{ij} - \mu \mathbf{1})^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}).$$

Calculando la derivada parcial respecto a μ :

$$\frac{\partial}{\partial \mu} [(Z_{ij} - \mu \mathbf{1})^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1})] = -2\mathbf{1}^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}).$$

Sustituyendo en la función objetivo, la derivada se convierte en:

$$\frac{\partial}{\partial \mu} \sum_{i=1}^n \sum_{j \neq i} -\frac{1}{2} (Z_{ij} - \mu \mathbf{1})^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}) w_{ij} = \sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}).$$

Igualando a cero para maximizar:

$$\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} (Z_{ij} - \mu \mathbf{1}) = 0.$$

Expandiendo el término:

$$\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} Z_{ij} - \mu \sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1} = 0.$$

Despejando μ :

$$\hat{\mu} = \frac{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} Z_{ij}}{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}}.$$

4.3 Propiedad de Inssegadez

Para demostrar que el estimador es inssegado, se calcula $E(\hat{\mu})$, la esperanza del estimador $\hat{\mu}$. El estimador está dado por:

$$\hat{\mu} = \frac{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} Z_{ij}}{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}}.$$

Tomando la esperanza, se tiene:

$$E(\hat{\mu}) = \frac{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} E(Z_{ij})}{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}}.$$

Se sabe que $Z_{ij} = \begin{pmatrix} Z(s_i) \\ Z(s_j) \end{pmatrix}$ tiene media $E(Z_{ij}) = \mu \mathbf{1}$, donde $\mathbf{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Sustituyendo esta relación, se obtiene:

$$E(\hat{\mu}) = \frac{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} (\mu \mathbf{1})}{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}}.$$

Expandiendo el numerador:

$$\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} (\mu \mathbf{1}) = \mu \sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}.$$

Por lo tanto:

$$E(\hat{\mu}) = \frac{\mu \sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}}{\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}}.$$

Obsérvese que el término $\sum_{i=1}^n \sum_{j \neq i} w_{ij} \mathbf{1}^\top \Sigma^{-1} \mathbf{1}$ aparece en el numerador y denominador, por lo que se simplifica a:

$$E(\hat{\mu}) = \mu.$$

Esto muestra que la esperanza del estimador coincide con el valor verdadero de μ , es decir, $\hat{\mu}$ es insesgado.

5 Pregunta 5

5.1 Parte 1: Predictor óptimo y error cuadrático medio

Se quiere demostrar que el error cuadrático medio:

$$EQM(g(Z)) = \mathbb{E}[(Z(s_0) - g(Z))^2]$$

es mínimo cuando $g(Z) = \mathbb{E}[Z(s_0) | Z]$, y que:

$$EQM(\mathbb{E}[Z(s_0) | Z]) = \text{Var}(Z(s_0) | Z).$$

El error cuadrático medio se define como:

$$EQM(g(Z)) = \mathbb{E}[Z(s_0)^2] - 2\mathbb{E}[Z(s_0)g(Z)] + \mathbb{E}[g(Z)^2].$$

Si $g(Z) = \mathbb{E}[Z(s_0) | Z]$, el segundo término $\mathbb{E}[Z(s_0)g(Z)]$ se maximiza debido a la propiedad de optimización de la media condicional:

$$\mathbb{E}[Z(s_0)g(Z)] = \mathbb{E}[Z(s_0)\mathbb{E}[Z(s_0) | Z]] = \mathbb{E}[\mathbb{E}[Z(s_0)^2 | Z]].$$

Además, $\mathbb{E}[g(Z)^2] = \mathbb{E}[\mathbb{E}[Z(s_0) | Z]^2]$. Por lo tanto, el error cuadrático medio queda como:

$$EQM(\mathbb{E}[Z(s_0) | Z]) = \mathbb{E}[Z(s_0)^2] - \mathbb{E}[\mathbb{E}[Z(s_0) | Z]^2].$$

Usando la propiedad de la varianza condicional:

$$\text{Var}(Z(s_0) | Z) = \mathbb{E}[Z(s_0)^2 | Z] - \mathbb{E}[\mathbb{E}[Z(s_0) | Z]^2 | Z],$$

se tiene que:

$$EQM(\mathbb{E}[Z(s_0) | Z]) = \text{Var}(Z(s_0) | Z).$$

Por lo tanto, el error cuadrático medio es mínimo cuando $g(Z) = \mathbb{E}[Z(s_0) | Z]$, y el valor mínimo del error cuadrático medio es igual a $\text{Var}(Z(s_0) | Z)$.

5.2 Parte 2: Predictor lineal

Sea $g(Z) = \lambda_0 + \lambda^\top Z$ un predictor lineal de $Z(s_0)$. Se quiere demostrar que si el predictor es insesgado, es decir, $\mathbb{E}[g(Z)] = \mathbb{E}[Z(s_0)] = \mu$, entonces:

$$EQM(g(Z)) = \mathbb{E}[(Z(s_0) - g(Z))^2]$$

es mínimo cuando:

$$g(Z) = \mu + r^\top R^{-1}(Z - \mathbf{1}\mu),$$

y que el error cuadrático medio asociado es:

$$EQM(\mu + r^\top R^{-1}(Z - \mathbf{1}\mu)) = \sigma^2(1 - r^\top R^{-1}r),$$

donde $r = [\rho(s_0 - s_i)]_{i=1}^n$ y $R = [\rho(s_i - s_j)]_{i,j=1}^n$.

5.2.1 Restricción de insesgadez

El predictor lineal es insesgado si:

$$\mathbb{E}[g(Z)] = \lambda_0 + \lambda^\top \mathbb{E}[Z] = \mu.$$

Dado que $\mathbb{E}[Z] = \mathbf{1}\mu$, se tiene:

$$\lambda_0 + \lambda^\top \mathbf{1}\mu = \mu.$$

Resolviendo para λ_0 :

$$\lambda_0 = \mu(1 - \lambda^\top \mathbf{1}).$$

5.2.2 Minimización del error cuadrático medio

El error cuadrático medio se define como:

$$EQM(g(Z)) = \mathbb{E}[(Z(s_0) - (\lambda_0 + \lambda^\top Z))^2].$$

Expandiendo:

$$EQM(g(Z)) = \mathbb{E}[Z(s_0)^2] - 2\mathbb{E}[Z(s_0)(\lambda_0 + \lambda^\top Z)] + \mathbb{E}[(\lambda_0 + \lambda^\top Z)^2].$$

Sustituyendo $\mathbb{E}[Z(s_0)] = \mu$, $\mathbb{E}[Z] = \mathbf{1}\mu$, y $\mathbb{E}[Z(s_0)Z] = r\sigma^2$, y usando la insesgadez de $g(Z)$, el error cuadrático medio queda como:

$$EQM(g(Z)) = \sigma^2 - 2\lambda^\top r\sigma^2 + \lambda^\top R\sigma^2\lambda.$$

Minimizando con respecto a λ , derivando:

$$\frac{\partial EQM}{\partial \lambda} = -2r\sigma^2 + 2R\sigma^2\lambda = 0.$$

Resolviendo, se obtiene:

$$\lambda = R^{-1}r.$$

5.2.3 Resultado final

Sustituyendo λ y λ_0 en $g(Z)$, se obtiene:

$$g(Z) = \mu + r^\top R^{-1}(Z - \mathbf{1}\mu).$$

El error cuadrático medio asociado es:

$$EQM = \sigma^2(1 - r^\top R^{-1}r).$$

6 Pregunta 6

Considérese el campo aleatorio estacionario de tipo Bernoulli $B = \{B(s)\}$, definido como:

$$B(s) = \begin{cases} 1 & \text{si } Z(s) < \mu(s), \\ 0 & \text{en otro caso (e.o.c.)}. \end{cases}$$

donde $Z(s)$ es un campo aleatorio Gaussiano de media 0, varianza 1 y función de correlación $\rho(h)$.

6.1 Distribución marginal del campo aleatorio B

La probabilidad marginal de $B(s)$ para $m = 0, 1$ está dada por:

$$\Pr(B(s) = m),$$

donde $B(s) = 1$ si $Z(s) < \mu(s)$ y $B(s) = 0$ en otro caso.

Para $B(s) = 1$, se cumple:

$$\Pr(B(s) = 1) = \Pr(Z(s) < \mu(s)).$$

Como $Z(s) \sim N(0, 1)$, esta probabilidad corresponde a la función de distribución acumulada de una variable normal estándar evaluada en $\mu(s)$:

$$\Pr(B(s) = 1) = \Phi(\mu(s)),$$

donde $\Phi(\cdot)$ es la función de distribución acumulada de la distribución normal estándar.

De manera similar, para $B(s) = 0$, se tiene:

$$\Pr(B(s) = 0) = 1 - \Phi(\mu(s)).$$

Por lo tanto, la distribución marginal del campo aleatorio $B(s)$ está comple-

tamente caracterizada por:

$$\Pr(B(s) = m) = \begin{cases} \Phi(\mu(s)), & \text{si } m = 1, \\ 1 - \Phi(\mu(s)), & \text{si } m = 0. \end{cases}$$

6.2 Distribución bivariada del campo aleatorio B

Ahora considérese la probabilidad bivariada:

$$\Pr(B(s_i) = m, B(s_j) = n), \quad m, n = 0, 1.$$

Dado que $B(s_i)$ y $B(s_j)$ están definidos a partir del campo aleatorio Gaussiano $Z(s)$, se debe considerar la distribución conjunta de $Z(s_i)$ y $Z(s_j)$. Se sabe que:

$$\begin{pmatrix} Z(s_i) \\ Z(s_j) \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho(h) \\ \rho(h) & 1 \end{pmatrix} \right),$$

donde $\rho(h)$ es la correlación entre $Z(s_i)$ y $Z(s_j)$ con $h = \|s_i - s_j\|$.

Para calcular $\Pr(B(s_i) = m, B(s_j) = n)$, considérense los siguientes casos:

Caso $m = 1, n = 1$:

$$\Pr(B(s_i) = 1, B(s_j) = 1) = \Pr(Z(s_i) < \mu(s_i), Z(s_j) < \mu(s_j)).$$

Esto corresponde a la probabilidad en la región $(-\infty, \mu(s_i)) \times (-\infty, \mu(s_j))$ de la distribución conjunta bivariada normal. Esta probabilidad se calcula como:

$$\Pr(B(s_i) = 1, B(s_j) = 1) = \Phi_2(\mu(s_i), \mu(s_j); \rho(h)),$$

donde $\Phi_2(\cdot, \cdot; \rho(h))$ es la función de distribución acumulada bivariada normal con correlación $\rho(h)$.

Caso $m = 1, n = 0$:

$$\Pr(B(s_i) = 1, B(s_j) = 0) = \Pr(Z(s_i) < \mu(s_i), Z(s_j) \geq \mu(s_j)).$$

Esto corresponde a:

$$\Pr(B(s_i) = 1, B(s_j) = 0) = \Phi(\mu(s_i)) - \Phi_2(\mu(s_i), \mu(s_j); \rho(h)).$$

Caso $m = 0, n = 1$:

$$\Pr(B(s_i) = 0, B(s_j) = 1) = \Pr(Z(s_i) \geq \mu(s_i), Z(s_j) < \mu(s_j)).$$

De manera similar, esta probabilidad es:

$$\Pr(B(s_i) = 0, B(s_j) = 1) = \Phi(\mu(s_j)) - \Phi_2(\mu(s_i), \mu(s_j); \rho(h)).$$

Caso $m = 0, n = 0$:

$$\Pr(B(s_i) = 0, B(s_j) = 0) = \Pr(Z(s_i) \geq \mu(s_i), Z(s_j) \geq \mu(s_j)).$$

Esto corresponde a:

$$\Pr(B(s_i) = 0, B(s_j) = 0) = 1 - \Phi(\mu(s_i)) - \Phi(\mu(s_j)) + \Phi_2(\mu(s_i), \mu(s_j); \rho(h)).$$

6.3 Función de covarianza del campo aleatorio B

La covarianza del campo aleatorio B está definida como:

$$\text{Cov}(B(s_i), B(s_j)) = \mathbb{E}[B(s_i)B(s_j)] - \mathbb{E}[B(s_i)]\mathbb{E}[B(s_j)].$$

Se sabe que:

$$\mathbb{E}[B(s_i)] = \Pr(B(s_i) = 1) = \Phi(\mu(s_i)),$$

$$\mathbb{E}[B(s_j)] = \Pr(B(s_j) = 1) = \Phi(\mu(s_j)).$$

Además:

$$\mathbb{E}[B(s_i)B(s_j)] = \Pr(B(s_i) = 1, B(s_j) = 1) = \Phi_2(\mu(s_i), \mu(s_j); \rho(h)).$$

Sustituyendo en la definición de la covarianza, se obtiene:

$$\text{Cov}(B(s_i), B(s_j)) = \Phi_2(\mu(s_i), \mu(s_j); \rho(h)) - \Phi(\mu(s_i))\Phi(\mu(s_j)).$$