7. Kriging con media desconocida. Kriging universal

Las situaciones más frecuentes en la práctica son aquellas en las que la media $m(x)$ es desconocida. Fijando ciertas hipótesis para esta función los sistemas de kriging ofrecen un predictor óptimo en el sentido que minimizan la varianza de predicción.

7.1. Variables estacionarias de segundo orden

7.1.1. Hipótesis iniciales

Dada la variable regionalizada $Z(x) = m(x) + Y(x)$, supongamos que:

- $E[Z(x)] = m(x) = \sum_{l=0}^{L} a_l f^l(x)$, donde f^l son funciones conocidas y a_l son parámetros reales desconocidos.

  Suele suponerse $f^0 = 1$, con lo que se asegura que el caso de media constante está incluido en el modelo.

- $Y(x)$ es estacionaria de segundo orden.

Predicción:

Dado un punto $x_0 \in D$ se desea construir el mejor predictor lineal insesgado (BLUE) de la variable $Z(x_0)$.

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) + \lambda_0.$$

1. Condición de insesgadez.

   Para que el predictor sea insesgado son condiciones necesarias $\lambda_0 = 0$ y $f^l(x_0) = \sum_{i=1}^{n} \lambda_i f^l(x_i)$; $l = 0, 1, \ldots, L$. Si ambas se verifican, entonces:

   $$E[Z^*(x_0)] = \sum_{l=0}^{L} a_l f^l(x_0) = E[Z(x_0)].$$

2. Condición de varianza mínima.

   Se trata de un problema de minimización de varianza de predicción sujeto a las $L+1$ restricciones necesarias para asegurar la insesgadez del predictor. La función a minimizar se plantea utilizando $L+1$ multiplicadores de Lagange μ_l:

$$Var\left(Z^*\left(x_0\right)-Z\left(x_0\right)\right)+2\sum_{l=0}^{L}\mu_l\left(\sum_{i=1}^{n}\lambda_i f^l\left(x_i\right)-f^l\left(x_0\right)\right).$$

El problema se reduce a encontrar un vector $X^{'}=(\lambda,\mu)^{'}=(\lambda_1,\lambda_2,\ldots,\lambda_n,$ $\mu_0,\mu_1,\ldots,\mu_L)$ que minimice la expresión anterior, la cual puede reescribirse de la siguiente forma:

$$\begin{aligned}
& Var\left(Z^*\left(x_0\right)-Z\left(x_0\right)\right)+2\sum_{l=0}^{L}\mu_l\left(\sum_{i=1}^{n}\lambda_i f^l\left(x_i\right)-f^l\left(x_0\right)\right)\\
= \; & E\left[\left(Z^*\left(x_0\right)-Z\left(x_0\right)\right)^2\right]+2\sum_{l=0}^{L}\mu_l\left(\sum_{i=1}^{n}\lambda_i f^l\left(x_i\right)-f^l\left(x_0\right)\right)\\
= \; & E\left[\left(\sum_{i=1}^{n}\lambda_i Y\left(x_i\right)-Y\left(x_0\right)\right)^2\right]+2\sum_{l=0}^{L}\mu_l\left(\sum_{i=1}^{n}\lambda_i f^l\left(x_i\right)-f^l\left(x_0\right)\right).
\end{aligned}$$

Calculando las derivadas parciales de esta última expresión con respecto a λ_i y a μ_l se obtiene el sistema matricial del cual se deduce el vector $X^{'}$. Si se desea ver el desarrollo de estas operaciones con más detalle puede consultarse alguno de los textos citados en la bibliografía.

7.1.2. Solución

El sistema matricial viene dado por:

$$\begin{pmatrix}
C(0) & C(x_1-x_2) & \cdots & C(x_1-x_n) & f^0(x_1) & \cdots & f^L(x_1)\\
C(x_2-x_1) & C(0) & \cdots & C(x_2-x_n) & f^0(x_2) & \cdots & f^L(x_2)\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
C(x_n-x_1) & C(x_n-x_2) & \cdots & C(0) & f^0(x_n) & \cdots & f^L(x_n)\\
f^0(x_1) & f^0(x_2) & \cdots & f^0(x_n) & 0 & \cdots & 0\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
f^L(x_1) & f^L(x_2) & \cdots & f^L(x_n) & 0 & \cdots & 0
\end{pmatrix}
\begin{pmatrix}
\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\\ \mu_0\\ \vdots\\ \mu_L
\end{pmatrix}$$

$$
= \begin{pmatrix} C(x_1 - x_0) \\ C(x_2 - x_0) \\ \vdots \\ C(x_n - x_0) \\ f^0(x_0) \\ \vdots \\ f^L(x_0) \end{pmatrix}
$$

$$
\Longleftrightarrow \begin{pmatrix} \Sigma & F \\ F' & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \sigma_0 \\ f_0 \end{pmatrix} \Longleftrightarrow AX = B.
$$

Siempre que A sea no singular, exite una única solución dada por

$$
X = A^{-1}B.
$$

Varianza de la predicción

Utilizando la expresión obtenida para X puede obtenerse el valor exacto de la varianza de la predicción.

$$
\begin{aligned}
\sigma^2_{KU} &= Var\left(Z^*(x_0) - Z(x_0)\right) = E\left[\left(\sum_{i=1}^{n} \lambda_i Y(x_i) - Y(x_0)\right)^2\right] \\
&= \sum_{i,j=1}^{n} \lambda_i \lambda_j C(x_i - x_j) - 2\sum_{i=1}^{n} \lambda_i C(x_i - x_0) + C(0)
\end{aligned}
$$

Puesto que $C(0) = \sigma^2_Z$ y que, tal como puede comprobarse en el sistema matricial,

$$
\sum_{j=1}^{n} \lambda_j C(x_i - x_j) = C(x_i - x_0) - \sum_{l=0}^{L} \mu_l f^l(x_i)
$$

se obtiene que

$$
\sigma^2_{KU} = \sigma^2_Z - \sum_{i=1}^{n} \lambda_i C(x_i - x_0) - \sum_{i=1}^{n} \lambda_i \sum_{l=0}^{L} \mu_l f^l(x_i)
$$

Finalmente, puesto que $\sum_{i=1}^{n} \lambda_i f^l(x_i) = f^l(x_0)$, podemos concluir que

$$
\sigma^2_{KU} = \sigma^2_Z - X'B.
$$

7.1.3. Propiedades

Propiedad de exactitud

Si $x_0 = x_i$ para algún $i = 1, 2, \ldots, n$, entonces $Z^*(x_0) = Z(x_0)$ y $\sigma^2_{KU} = 0$. Es decir, el sistema de kriging es un interpolador.

7.2. Variables intrínsecas

Si el proceso $Z(x)$ no es estacionario de segundo orden, pero por lo menos verifica la hipótesis intrínseca pueden plantearse los sistemas de ecuaciones del kriging para obtener un predictor lineal puntual. Aunque los desarrollos son muy similares ahora resultan un poco más engorrosos en algunos de sus pasos. Como podréis comprobar al final de esta sección, los resultados a los que se llegan son parecidos al caso de estacionariedad de segundo orden.

7.2.1. Hipótesis iniciales

- $E[Z(x)] = \sum_{l=0}^{L} a_l f^l(x)$, donde f^l, son funciones conocidas y a_l son parámetros reales desconocidos.
- $Y(x)$ verifica la hipótesis intrínseca.

Predicción:

Dado un punto $x_0 \in D$ se desea construir el mejor predictor lineal insesgado (BLUE) de la variable $Z(x_0)$.

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) + k.$$

Las condiciones de insesgadez y de varianza mínima implican que $k = 0$ y que debemos encontrar un vector $X^{'} = (\lambda, \mu)^{'} = (\lambda_1, \lambda_2, \ldots, \lambda_n, \mu_0, \mu_1, \ldots, \mu_L)$ que minimice

$$\begin{aligned}
& Var\left(Z^*\left(x_0\right) - Z\left(x_0\right)\right) - 2\sum_{l=0}^{L} \mu_l \left(\sum_{i=1}^{n} \lambda_i f^l\left(x_i\right) - f^l\left(x_0\right)\right) \\
= \ & E\left[\left(Z^*\left(x_0\right) - Z\left(x_0\right)\right)^2\right] - 2\sum_{l=0}^{L} \mu_l \left(\sum_{i=1}^{n} \lambda_i f^l\left(x_i\right) - f^l\left(x_0\right)\right) \\
= \ & -\sum_{i=1}^{n}\sum_{l=1}^{n} \lambda_i \lambda_l \gamma\left(x_i - x_l\right) + 2\sum_{i=1}^{n} \lambda_i \gamma\left(x_i - x_0\right) \\
& -2\sum_{l=0}^{L} \mu_l \left(\sum_{i=1}^{n} \lambda_i f^l\left(x_i\right) - f^l\left(x_0\right)\right).
\end{aligned}$$

Las derivadas parciales de esta última expresión dan lugar a un sistema matricial del cual se deduce el vector $X^{'}$. Si se desea ver el desarrollo de estas operaciones con más detalle puede consultarse alguno de los textos citados en la bibliografía.

7.2.2. Solución

El sistema matricial viene dado por:

$$\begin{pmatrix}
\gamma(0) & \gamma(x_1 - x_2) & \cdots & \gamma(x_1 - x_n) & f^0(x_1) & \cdots & f^L(x_1) \\
\gamma(x_2 - x_1) & \gamma(0) & \cdots & \gamma(x_2 - x_n) & f^0(x_2) & \cdots & f^L(x_2) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\gamma(x_n - x_1) & \gamma(x_n - x_2) & \cdots & \gamma(0) & f^0(x_n) & \cdots & f^L(x_n) \\
f^0(x_1) & f^0(x_2) & \cdots & f^0(x_n) & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
f^L(x_1) & f^L(x_2) & \cdots & f^L(x_n) & 0 & \cdots & 0
\end{pmatrix}
\begin{pmatrix}
\lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu_0 \\ \vdots \\ \mu_L
\end{pmatrix}$$

$$= \begin{pmatrix} \gamma(x_1 - x_0) \\ \gamma(x_2 - x_0) \\ \vdots \\ \gamma(x_n - x_0) \\ f^0(x_0) \\ \vdots \\ f^L(x_0) \end{pmatrix}$$

$$\Longleftrightarrow \begin{pmatrix} \Gamma & F \\ F' & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma_0 \\ f_0 \end{pmatrix} \Longleftrightarrow AX = B.$$

Siempre que A sea no singular, exite una única solución dada por

$$X = A^{-1}B.$$

Varianza de la predicción

Utilizando la expresión de X puede obtenerse el valor exacto de la varianza del predicción.

$$\begin{aligned} \sigma^2_{KU} &= Var\left(Z^*(x_0) - Z(x_0)\right) = E\left[\left(Z^*(x_0) - Z(x_0)\right)^2\right] \\ &= X^{'}B. \end{aligned}$$

7.2.3. Propiedades

Si $x_0 = x_i$ para algún $i = 1, 2, \ldots, n$, entonces $Z^*(x_0) = Z(x_0)$ y $\sigma^2_{KU} = 0$. Es decir, el sistema de kriging es un interpolador.

7.3. Kriging con media desconocida pero constante. Kriging ordinario

Este es un caso particular del kriging universal, en el cual $L = 0$. Por esta razón obviaremos los desarrollos, pasando directamente a mostrar las expresiones necesarias para obtener la predicción y la varianza kriging.

7.3.1. Variables estacionarias de segundo orden

Hipótesis iniciales

- $E[Z(x)] = m$, $\forall x \in D$, con m constante desconocida.
- $Y(x)$ es estacionaria de segundo orden.

Predicción y solución

Bajo estas condiciones el sistema matricial a resolver es:

$$\begin{pmatrix} C(0) & C(x_1 - x_2) & \cdots & C(x_1 - x_n) & 1 \\ C(x_2 - x_1) & C(0) & \cdots & C(x_2 - x_n) & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ C(x_n - x_1) & C(x_n - x_2) & \cdots & C(0) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix}$$

$$= \begin{pmatrix} C(x_1 - x_0) \\ C(x_2 - x_0) \\ \vdots \\ C(x_n - x_0) \\ 1 \end{pmatrix}$$

$$\Longleftrightarrow \begin{pmatrix} \Sigma & F \\ F' & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \sigma_0 \\ f_0 \end{pmatrix} \Longleftrightarrow AX = B$$

Siempre que A sea estrictamente definida positiva, exite una única solución dada por

$$X = A^{-1}B.$$

Varianza de la predicción

$$\begin{aligned} \sigma^2_{KO} &= Var\left(Z^*(x_0) - Z(x_0)\right) = E\left[\left(Z^*(x_0) - Z(x_0)\right)^2\right] \\ &= \sigma^2_Z - X'B. \end{aligned}$$

7.3.2. Variables intrínsecas

Hipótesis iniciales

- $E[Z(x)] = m$, $\forall x \in D$, con m constante desconocida.

- $Y(x)$ verifica la hipótesis intrínseca

Predicción y solución

En este caso el sistema matricial viene dado por

$$\begin{pmatrix} \gamma(0) & \gamma(x_1-x_2) & \cdots & \gamma(x_1-x_n) & 1 \\ \gamma(x_2-x_1) & \gamma(0) & \cdots & \gamma(x_2-x_n) & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma(x_n-x_1) & \gamma(x_n-x_2) & \cdots & \gamma(0) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix}$$

$$= \begin{pmatrix} \gamma(x_1-x_0) \\ \gamma(x_2-x_0) \\ \vdots \\ \gamma(x_n-x_0) \\ 1 \end{pmatrix}$$

$$\Longleftrightarrow \begin{pmatrix} \Gamma & F \\ F' & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma_0 \\ f_0 \end{pmatrix} \Longleftrightarrow AX = B.$$

Siempre que A sea estrictamente definida negativa, exite una única solución dada por

$$X = A^{-1}B.$$

Varianza de la predicción

$$\begin{aligned} \sigma_{KO}^2 &= Var\left(Z^*(x_0) - Z(x_0)\right) = E\left[\left(Z^*(x_0) - Z(x_0)\right)^2\right] \\ &= X'B. \end{aligned}$$

7.4. Observaciones sobre los sistemas de kriging

- En kriging ordinario y universal se verifica que $\sum_{i=1}^n \lambda_i = 1$. En kriging simple no tiene por qué verificarse esta igualdad.

- En caso de no conocerse las verdaderas expresiones de $C(h)$ y $\gamma(h)$, éstas deben sustituirse por sus estimaciones $\hat{C}(h)$ y $\hat{\gamma}(h)$ en los sistemas de kriging. Esto quiere decir que el error final de la predicción será algo mayor que el que nos indique la varianza de la predicción.

- Bajo las hipótesis del kriging universal, el vector de parámetros $a^{'} = (a_0, a_2, \ldots, a_L)$ debería ser estimado mediante mínimos cuadrados generalizados, es decir, determinar el a que minimice

$$(Z - Fa)^{'} \Sigma^{-1} (Z - Fa) = Y^{'} \Sigma^{-1} Y,$$

donde

$$\begin{aligned}
\Sigma &= \begin{pmatrix} Cov(Z(x_1), Z(x_1)) & \cdots & Cov(Z(x_1), Z(x_n)) \\ \cdots & \cdots & \cdots \\ Cov(Z(x_n), Z(x_1)) & \cdots & Cov(Z(x_n), Z(x_n)) \end{pmatrix}, \\
F &= \begin{pmatrix} f^0(x_1) & \cdots & f^L(x_1) \\ \cdots & \cdots & \cdots \\ f^0(x_n) & \cdots & f^L(x_n) \end{pmatrix}, \\
Z^{'} &= (Z(x_1), Z(x_2), \ldots, Z(x_n)), \\
Y^{'} &= (Y(x_1), Y(x_2), \ldots, Y(x_n)).
\end{aligned}$$

La solución viene dada por

$$\widehat{a}_{MCG} = \left(F^{'} \Sigma^{-1} F\right)^{-1} F^{'} \Sigma^{-1} F.$$

- En kriging universal $Y(x)$ es la variable que presenta la estructura de dependencia espacial con la cual debe estimarse $C(h)$ o $\gamma(h)$. Por otra parte $C(h)$ o $\gamma(h)$ son necesarias para predecir el vector a y poder así obtener $Y(x)$.

  Para solucionar este problema, Cressie (1993) propone el siguiente algoritmo:

  1. Primero se estima a mediante $\widehat{a}_{MCO}$.
  2. Se obtiene $\widehat{m}(x) = \sum_{l=0}^{L} \widehat{a}_l f^l(x)$.
  3. Se estima $\widehat{Y}(x) = Z(x) - \widehat{m}(x)$ y $\widehat{\Sigma}$.

4. Se estima a mediante $\widehat{a}^1_{MCG} = \left(F'\widehat{\Sigma}^{-1}F\right)^{-1} F'\widehat{\Sigma}^{-1}Z$.
5. Se repite 2-3-4 hasta que $\left\|\widehat{a}^i_{MCG} - \widehat{a}^{i+1}_{MCG}\right\| < \epsilon$, con ϵ un error determinado por el usuario.

El hecho de trabajar con $\widehat{Y}$ en vez de con Y introducirá sesgo en la estimación de $C(h)$ y/o $\gamma(h)$.

- Bajo las hipótesis de kriging ordinario y con $Y(x)$ variable estacionaria de segundo orden:

  1. Si $\{Z(x_1), Z(x_2), \ldots, Z(x_n)\}$ incorreladas dos a dos:
     $\lambda_i = \rho_{i0} - \mu/\sigma^2_Z; \quad i = 1, 2, \ldots, n.$
     $\mu = \left(\sum_{j=1}^n C_{j0} - \sigma^2_Z\right)/n \Longrightarrow \lambda_i = \rho_{i0} + \left(1 - \sum_{j=1}^n \rho_{j0}\right)/n;\ i = 1, 2, \ldots, n.$
     $\sigma^2_{KO} = \sigma^2_Z - \sum_{i=1}^n C_{i0}\rho_{i0} + \mu\left(\sum_{i=1}^n \rho_{i0} - 1\right).$
  2. Si $\{Z(x_1), Z(x_2), \ldots, Z(x_n), Z(x_0)\}$ incorreladas dos a dos:
     $\lambda_i = 1/n; \quad i = 1, 2, \ldots, n.$
     $\mu = -\sigma^2_Z/n$
     $\sigma^2_{KO} = \sigma^2_Z + \sigma^2_Z/n.$
  3. De las expresiones anteriores se deduce que el kriging ordinario, en el caso de datos incorrelados, resuelve la predicción de $Z(x_0)$ mediante técnicas de regresión:
     $$Z^*(x_0) = \frac{1}{n}\sum_{i=1}^n Z(x_i) = \hat{a}_0 = \hat{m}$$
     y que la varianza de predicción está acotada superiormente si la variable es estacionaria de segundo orden.

- Bajo las hipótesis de kriging universal y con $Y(x)$ variable estacionaria de segundo orden:

  1. Si $\{Z(x_1), Z(x_2), \ldots, Z(x_n)\}$ incorreladas dos a dos:
     $\lambda_i = \rho_{i0} - \sum_{l=0}^L f^l(x_i)\mu_l/\sigma^2_Z; \quad i = 1, 2, \ldots, n.$
     $\sigma^2_{KU} = \sigma^2_Z - \sum_{i=1}^n C_{i0}\rho_{i0} + \sum_{l=0}^L \mu_l\left[\sum_{i=1}^n f^l(x_i)\rho_{i0} - f^l(x_0)\right].$

Los multiplicadores de Lagrange deben calcularse según:

$$\begin{pmatrix} \sum_{k=1}^{n} f^0(x_k) f^0(x_k) & \cdots & \sum_{k=1}^{n} f^0(x_k) f^L(x_k) \\ \cdots & \cdots & \cdots \\ \sum_{k=1}^{n} f^L(x_k) f^0(x_k) & \cdots & \sum_{k=1}^{n} f^L(x_k) f^L(x_k) \end{pmatrix} \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_L \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{n} f^0(x_k) C_{k0} - \sigma_Z^2 f^0(x_0) \\ \vdots \\ \sum_{k=1}^{n} f^L(x_k) C_{k0} - \sigma_Z^2 f^L(x_0) \end{pmatrix}.$$

2. Si $\{Z(x_1), Z(x_2), \ldots, Z(x_n), Z(x_0)\}$ incorreladas dos a dos:

   $\lambda_i = -\sum_{l=0}^{L} f^l(x_i) \mu_l / \sigma_Z^2; \quad i = 1, 2, \ldots, n.$

   $\sigma_{KU}^2 = \sigma_Z^2 - \sum_{l=0}^{L} \mu_l f^l(x_0).$

   Los multiplicadores de Lagrange deben calcularse según:

$$\begin{pmatrix} \sum_{k=1}^{n} f^0(x_k) f^0(x_k) & \cdots & \sum_{k=1}^{n} f^0(x_k) f^L(x_k) \\ \cdots & \cdots & \cdots \\ \sum_{k=1}^{n} f^L(x_k) f^0(x_k) & \cdots & \sum_{k=1}^{n} f^L(x_k) f^L(x_k) \end{pmatrix} \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_L \end{pmatrix}$$

$$= \begin{pmatrix} -\sigma_Z^2 f^0(x_0) \\ \vdots \\ -\sigma_Z^2 f^L(x_0) \end{pmatrix}.$$

3. De las expresiones anteriores se deduce que el kriging universal, en el caso de datos incorrelados, resuelve la predicción de $Z(x_0)$ mediante técnicas de regresión:

$$Z^*(x_0) = \sum_{l=0}^{L} \hat{a}_l f^l(x) = \hat{m}(x)$$

   y que la varianza de predicción está acotada superiormente si la variable es estacionaria de segundo orden.