AKURASI METODE LAGRANGE DAN ORDINARY KRIGING UNTUK INTERPOLASI DATA MULTIVARIAT

Hery Andi Sitompul, S.Si, M.Si
Dosen Kopertis Wilayah I Sumut, Dpk Universitas Darma Agung, Medan.

Abstract
Interpolate for function of several variables or multivariate data is still very limited discussed in literature. Would be very useful if help who compares the result of both scheme interpolate whether deterministic and stochastic. Lagrange method is one of the deterministic interpolaton scheme who determine a polynomial function to approach sample data, while Ordinary Kriging is one of the scheme stochastic interpolation who use the parameter of statistic sample data in neighbor point that will be interpolate. Both this scheme have the same result that very fine in interpolate points on multivariate data, although both of this method need different sample size.
Keywords : Interpolation, Deterministic, Stochastic, Kriging, Lagrange.

PENDAHULUAN

LATAR BELAKANG

Sering kali kita dihadapkan pada permasalahan nilai sebuah fungsi $f(x)$ yang diketahui pada satu himpunan titik $x_1, x_2, \dots, x_N$ (katakanlah $x_1 < x_2 < \dots < x_N$) tetapi kita tidak mengetahui ekspresi analitik untuk $f(x)$ tersebut ,sehingga kita tidak bias menentukan nilai $f(x)$ pada sembarang titik x yang lain.Sebagai contoh, misalkan $f(x_i)$ merupakan hasil pengukuran fisik atau hasil perhitungan secara numerik, tetapi tidak dapat dinyatakan dalam bentuk fungsi. Sekarang jika kita ingin mengetahui nilai $f(x)$ pada sembarang titik yang berada diantara nilai terkecil dan terbesar x dapat dilakukan dengan estimasi. Salah satu metode yang sering digunakan adalah proses interpolasi.

Berbagai skema interpolasi untuk proses estimasi nilai $f(x)$/ fungsi variabel tunggal seperti yang dijelaskan sebelumnya sudah sangat banyak dibahas dan dipelajari. Jika diperhatikan dalam berbagai buku literature sudah sangat banyak membahas mengenai teori interpolasi untuk fungsi satu variable tetapi masih terbatas untuk fungsi polynomial multivariat. Padahal sangat banyak kasus dapat kita jumpai yang melibatkan data multivariate. Sebagai contoh adalah data geostatsistik, hidrologi, meteorologi, maupun interpolasi permukaan.

Terdapat dua bagian besar metode interpolasi yang melibatkan data multivariate yaitu secara deterministic dan stokastik.

Akan sangat bermanfaat dan memudahkan bagi pengguna berikutnya jika ada kajian yang membandingkan antara kedua metode ini.

Dalam tulisan ini metode deterministik yang dipilih adalah metode Lagrange multivariate dan metode stokastik adalah Ordinary Kriging

MAKSUD DAN TUJUAN

Maksud dari tulisan ini adalah untuk mempelajari perbedaan yang dihasilkan oleh metode interpolasi Lagrange Multivariat dan Ordinary Kriging.

Sedangkan tujuan dari tulisan ini adalah :

1. Mempelajari keuntungan dan kerugian dari interpolasi dengan menggunakan metode Lagrange Multivariat dan Ordinary Kriging
2. Menjelaskan Perbedaan yang dihasilkan
3. Menentukan metode yang terbaik untuk data multivariat.

LANDASAN TEORI

INTERPOLASI

Interpolasi adalah proses pencarian atau perhitungan suatu nilai fungsi pada sembarang variable dengan menggunakan data beberapa titik yang dilalui oleh fungsi tersebut.

Secara formal defenisi dari interpolasi adalah sebagai berikut :

Diketahui :

1. $n+1$daftar titik, $x_0, x_1, x_2, \ldots, x_N$
2. Nilai fungsi$f_0, f_1, f_2, \ldots, f_N$ pada titik tersebut
3. Nilai x pada $x_0 < x_I < x_N$ yang tidak berada dalam daftar

Prediksi/hitung : nilaif_I pada x_I

Interpolasi multivariate adalah adaptasi dari proses interpolasi fungsi satu variable untuk fungsi beberapa peubah. Atau disebut juga interpolasipermukaan.

LAGRANGE MULTIVARIAT

Metode Lagrange merupakan metode deterministic yang menggunakan bentuk polynomial dari data yang diberikan untuk menghitung suatu nilai fungsi.Bentuk umum interpolasi polynomial Lagrange adalah :

$$L_k(x) = \prod_{j=1, j\neq k}^{n} \frac{x - x_j}{x_k - x_j}$$

Untuk data yang diberikan :

x	x_1	x_2	x_3	.	.	.	x_n
$f(x)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	.	.	.	$f(x_n)$

Bentuk polynomial untuk kasus multivariate adalah :

Misalkan $f = f(X_1, X_2, X_3, \ldots, X_n)$ adalah sebuah fungsi dengan n variable bebas yang berderajat m, oleh karena itu akan terdapat $\binom{m+n}{m} = \rho$ bentuk polynomial yang mungkin dalam f , dengan demikian akan terdapat sebanyak ρ data yang diketahui $(x_{1i}, x_{2i}, \ldots, x_{ni}, f_i) \in R^{n+1}$,$1 \leq i \leq \rho$, dengan $f_i = f(x_{1i}, x_{2i}, \ldots, x_{ni})$ atau dalam bentuk lain :

$f(X_1, X_2, X_3, \ldots, X_n) = \sum_{e_{i.1 \leq m}} \alpha_{e_i} X^{e_i}$dimana$\alpha_{e_i}$ adalah koefisien dalam f, $X = (X_1, X_2, X_3, \ldots, X_n)$ adalah n buah variable bebas dari f, $e_i = (e_1, e_2, \ldots, e_n)$ adalah vektor eksponen dengan elemen integer non negatif antara 1 dan m. Dengan demikian $X^{e_i} = \prod_{j=1}^{n} X_j^{e_i}$. Mengikuti bentuk penulisan polynomial Lagrange, f dapat ditulis dalam bentuk $\sum_{i1}^{\rho} f_i l_i(X)$, dimana$l_i(X)$ adalah fungsi multinomial dengan variabel bebas $(X_1, X_2, X_3, \ldots, X_n)$ dengan sifat jika $X = (x_{1i}, x_{2i}, \ldots, x_{ni})$ maka $l_i(X_i) = 1$ dan $l_i(X_i) = 0 (j \neq i)$.

Dari system persamaan linier $f_i = \sum_{e_{i.1 \leq m}} \alpha_{e_i} X^{e_i}$ dimana$1 \leq i \leq \rho$ dapat dikonstruksikan matriks sampel= $\left[X_j^{e_i}\right]$:

$$M = \left[\begin{pmatrix} X_1^{e_1} & \cdots & X_1^{e_\rho} \\ \vdots & \ddots & \vdots \\ X_\rho^{e_1} & \cdots & X_\rho^{e_\rho} \end{pmatrix}\right]$$

Misalkan $\Delta =$ det(M), maka multinomial untuk $l_i(X)$ adalah :

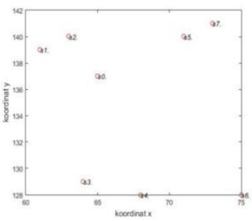
$l_i(X) = \frac{\Delta_i(X)}{\Delta}$dan

$$f = \sum_{i1}^{\rho} f_i \frac{\Delta_i(X)}{\Delta}$$

ORDINARY KRIGING

Kriging adalah metode geostatistik yang digunakan untuk mengestimasi nilai dari sebuah titik atau blok sebagai kombinasi linier dari nilai contoh yang terdapat disekitar titik yang akan diestimasi. Bobot kriging diperoleh dari hasil variansi estimasi minimum dengan memperluas penggunaan semi-variogram. Estimator kriging dapat diartikan sebagai variable tidak bias dan penjumlahan dari keseluruhan bobot adalah satu. Bobot inilah yang dipakai untuk mengestimasi nilai dari ketebalan, ketinggian, kadar atau variable lain (*https :/ /id. wikipedia. org/wiki/Kriging*).

Sebagai contoh diberikan dalam gambar berikut, titik yang akan diestimasi adalah s_0, dengan menggunakan data s_1,s_2,s_3,s_4,s_5,s_6,s_7 yang ada disekeliling s_0.



Gambar 1.Polai nterpolasi dengan Kriging.

Terdapat berbagai macam metode kriging, antara lain adalah : Simple Kriging, Universal Kriging, Ordinary Kriging, Co-Kriging, dll. Ordinary Kriging dipilih karena metode ini adalah yang paling umum /asli dan metode kriging yang lain dapat dimodifikasi dari metode ini.

Ordinary Kriging didasarkan pada beberapa model geostatistik yang selama ini sudah lama digunakan dengan batasan sebagai berikut :

1. Mean/rata-rata $\mu(s)$ diasumsikan adalah adalah konstan.
2. Variogram$\gamma(h)$ diasumsikan diketahui.

Misalkan s_0 adalah sebuah titik sembarang yang berada pada daerah D yang bukan sampel, dan akan diprediksi nilainya $(Z(s_0))$.

Sifat – sifat dari prediktor Ordinary Kriging :

1. Prediktor merupakan kombinasi linier dari data sampel$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i)$
2. Tidak bias, dan memenuhi : $E\left(\hat{Z}(s_0)\right) = E(Z(s_0))$
3. $\sum_{i=1}^{n} \lambda_i = 1$

Untuk mendapatkan predictor tak bias terbaik dapat dilakukan dengan meminimasi $var[\hat{Z}(s_0) - Z(s_0)]$ terhadap kendala $\sum_{i=1}^{n} \lambda_i = 1$. Dengan mengaplikasikan metode pengali lagrange dari kalkulus dapat diperoleh bahwa koefisien – koefisen $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ yang optimal adalah yang memenuhi persamaan :

$$\Gamma_O \lambda_O = \gamma_O$$

Dimana :

λ_O=$(\lambda_1, \lambda_2, \ldots, \lambda_n, \varrho)$

$\gamma_0 = [\gamma(s_1 - s_0), \gamma(s_2 - s_0), \ldots, \gamma(s_n - s_0)]'$adalah variogram

$$\Gamma_O = \begin{cases} \gamma(s_i - s_j) & i = 1, \ldots, n\,; j = 1, \ldots, n \\ 1, & i = n+1; j = 1,2, \ldots, n \\ 0, & i = n+1; j = n+1 \end{cases}$$

ϱ adalah pengali Lagrange dan Γ_O adalah matriks simetris.

Variansi minimum disebut variansi Kriging :

$$\sigma_{OK}^2(s_0) = \sum_{i=1}^{n} \lambda_i\, \gamma(s_i - s_0) + \varrho = \lambda_O{}'\gamma_O$$

Dalam kasus variogram $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n, \varrho)$ dapat diperoleh bahwa : $\lambda' = \Gamma'\gamma$

Dimana :

$$\Gamma = \begin{bmatrix} \gamma(1,1) & \gamma(1,2) & \cdots & \gamma(1,n) & 1 \\ \gamma(2,1) & \gamma(2,2) & \cdots & \gamma(2,n) & 1 \\ \vdots & \vdots & \vdots & \vdots & 1 \\ \gamma(n,1) & \gamma(n,2) & \cdots & \gamma(n,n) & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\gamma = \begin{bmatrix} \gamma(0,1) \\ \gamma(0,2) \\ \vdots \\ \gamma(0,n) \\ 1 \end{bmatrix}$$

Denganγ adalah variogram eksperimental. Model variogram ini dapat diperoleh dari data yang ada dengan menggunakan persamaan :

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n} [z(x_i + h) - z(x_i)]^2$$

x_i= lokasi titik sampel, $z(x_i)$= nilai dari lokasi titik sampel, $n\,(h)$= jumlah pasangan titik $(x_i, x_i + h)$ yang dipisahkan jarak h.

Selanjutnya untuk keperluan analisis variogram eksperimen tersebut harus diganti dengan model variogram teoritis yang paling sesuai, model variogram teoritis tersebut adalah sebagai berikut :

1. Model Spherical : $\gamma(h) = \begin{cases} C_1\left(\frac{3}{2}\frac{h}{a} - \frac{1}{2}\frac{h^3}{a^3}\right) & h \le a \\ C_1 & h > a \end{cases}$
2. Model Eksponensial : $\gamma(h) = C_1[1 - \exp\left(-\frac{h}{a}\right)]$
3. Model Gaussian : $\gamma(h) = C_1[1 - \exp\left(-\frac{h}{a}\right)^2]$
4. Model linier : $\gamma(h) = C_1|h|$

Dengan a= Range, h=jarak, C_1= sill.

METODOLOGI

Untuk memperlihatkan perbedaan akurasi yang dihasilkan oleh kedua metode ini, akan dilakukan dengan menerapkan kedua metode tersebut pada sebuah kasus yang sama, kemudian membandingkan hasil yang diperoleh. Proses penentuan model estimasi dilakukan dengan pemograman Matlab dan Maple

Kasus yang akan diterapkan pada kedua metode adalah : akan diinterpolasi berbagai titik yang dilalui oleh permukaan sebuah fungsi dengan ekspresi sebagai berikut:

$f: X \to f(X), X \in \Omega \subseteq \mathbb{R}^s, f(X) \in \mathbb{R}$

$\Omega = [-4,4]^2, X = (x, y)$

$f(X) = x^2 - y^2$ ....(*)

Interpolasi Data dengan Metode Langrange Multivariat

Beberapa titik yang dilalui oleh fungsi dalam persamaan (*) diatas diambil secara acak sebagai sampel yaitu sebagai berikut:

Tabel 1.Data titiks ampel.

No	x	y	$f(X)$
1	-2	-1	3
2	-1	2	-2
3	3	2	5
4	1	3	-8
5	2	-2	0
6	0	-3	-9

Dengan menggunakan metode interpolasi Lagrange Multivariat dapat diperoleh bahwa :

$\rho = 6, n = 2$ (jumlah variable bebas)

$$\binom{m+2}{m} = 6$$

$m = 2$ (orde polynomial)

Dengan demikian data sampel tersebut berada pada fungsi dengan dua peubah bebas dan berderajat 2, yang memenuhi system persamaan :

$$f_i = \sum_{e_{i.1 \le m}} \alpha_{e_i} X^{e_i}$$

$$f_i(x, y) = \alpha_1 x_i^2 + \alpha_1 x_i y_i + \alpha_3 y_i^2 + \alpha_4 x_i + \alpha_5 y_i + \alpha_6$$

Dengan mensubstitusi data sampel kedalam system persamaan diatas diperoleh matriks $M_{6\times6}$ yang berisiko efisien – koefisien $\alpha_1, \alpha_2, \ldots, \alpha_5, \alpha_6$. Sedemikian hingga :

$$\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_6 \end{bmatrix} = M \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_6 \end{bmatrix}$$

$$M = \begin{bmatrix} 4 & 2 & 1 & -2 & -1 & 1 \\ 1 & -2 & 4 & -1 & 2 & 1 \\ 9 & 6 & 4 & 3 & 2 & 1 \\ 1 & 3 & 9 & 1 & 3 & 1 \\ 4 & -4 & 4 & 2 & -2 & 1 \\ 0 & 0 & 9 & 0 & -3 & 1 \end{bmatrix}$$

$\Delta =$ $\det(M)$ $=1144$

M_i= Matriks dimana setiap baris ke-i pada M diganti dengan $x^2 \quad xy \quad y^2 \quad x \quad y \quad 1$

$\Delta_i = \det(M_i)$

maka multinomial untuk $l_i(X) = \frac{\Delta_i(X)}{\Delta}$adalah:

$$l_1(X) = -\frac{10}{13}x^2 + \frac{5}{13}xy - \frac{17}{26}y^2 + \frac{10}{13}x - \frac{5}{26}y + \frac{69}{13}$$

$$l_2(X) = \frac{673}{572}x^2 - \frac{15}{26}xy + \frac{509}{572}y^2 - \frac{829}{572}x + \frac{191}{572}y - \frac{1002}{143}$$

Demikian seterusnya dilakukan hingga $l_6(X)$.

$$l_6(X) = \frac{114}{143}x^2 - \frac{4}{13}xy + \frac{199}{286}y^2 - \frac{140}{143}x + \frac{5}{286}y - \frac{745}{143}$$

Setelah $l_1(X), l_2(X), \ldots, l_6(X)$ didapatkan maka :

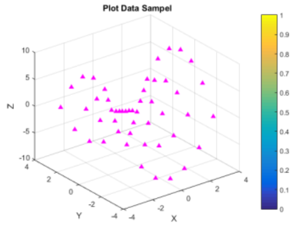
$$f = \sum_{i1}^{\rho} f_i \frac{\Delta_i(X)}{\Delta}$$

$$f(X) = l_1 + l_2 + \cdots + l_6 = x^2 - y^2 \ldots (**)$$

Interpolasi Data dengan Ordinary Kriging

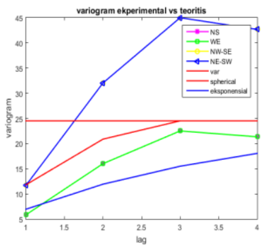
Estimasi dengan Ordinary Kriging dilakukan dengan analisis structural yaitu pencocokan variogram eksperimen dengan teoritis. Variogram eksperimental diperoleh dengan cara menghitung variansi data pengukuran/eksperimen untuk masing - masing lag jarak dan arah, selanjutnya nilai variansi tersebut dicocokkan dengan model variogram teoritis yang

ada sedemikian hingga diperoleh sebuah model variogram yang paling sesuai.
Data eksperimen yang digunakan dalam metode ini diambil secara acak dari fungsi pada persamaan (*) dan disajikan dalam plot berikut :



Gambar 2. Data titik sampel pada Ordinary Kriging.

Dari hasil analisa structural pada data sampel dan pencocokan variogram eksperimen dengan teoritis diperoleh bahwa model variogram terbaik adalah model spherical dengan range $(a = 3)$, sill $(C_1 = 24{,}5)$seperti terlihat dalam grafik berikut :



Gambar 3. Pencocokan variogram empiris dengan teoritis.

Maka diperoleh :

$$\gamma(h) = \begin{cases} C_1\left(\frac{3}{2}\frac{h}{a} - \frac{h^3}{2a^3}\right) & , h \le a \\ C_1 & h > a \end{cases} \quad ......(***)$$

Hasil dan Pembahasan

Model estimasi yang dihasilkan dari sampel pada kedua metode adalah sebagai berikut :

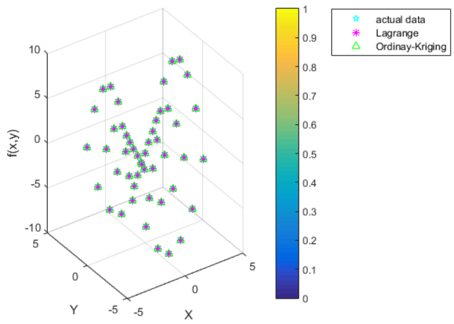
1. Metode Lagrange menghasilkan polynomial untuk interpolas iadalah : $f(x,y) = x^2 - y^2$
2. Metode Ordinary Kriging menghasilkan :$s_0 = \sum_{i=1}^{n} \lambda_i s_i$, dimana λ_i adalah bobot yang diperoleh dari $\lambda = \Gamma^{-1}\gamma$ dengan γ adalah persamaan (***).

Kedua model tersebut akan digunakan untuk menginterpolasi beberapa titik yang disajikan dalam table berikut :

Tabel 2.Hasil interpolasi di beberapa titik.

x	y	Lagrange	Ordinary Kriging	Data Aktual
-3	1	8	8.0000	8
-3	0	9	9.0000	9
1	-2	-3	-3.0000	-3
2	1	3	3.0000	3
0	-1	-1	-1.0000	-1
2	2	0	-8.9540e-14	0

Secara keseluruhan hasil interpolasi dari kedua model yang dihasilkan dapat dilihat dalam grafik berikut :



Gambar 4. Hasil interpolasi dengan kedua metode dibandingkan dengan data sebenarnya.

Tabel 3.Statistik hasil interpolasi

Metode	Total Error	MSE
Lagrange	0	0
Ordinary Kriging	-1.0969e-12	1.9417e-27

Dari hasil perhitungan yang disajikan dalam gambar 4 dan tabel 3 dapat terlihat bahwa kedua metode menghasilkan model estimasi sangat baik,

Kesimpulan

Beberapa hal yang dapat disimpulkan dari pembahasan diatas adalah :

1. Metode Lagrange dapat menghasilkan interpolasi yang sangat presisi, ini dikarenakan polynomial interpolasi yang

dihasilkan sama dengan fungsi yang melalui titik sampel.
2. Metode ordinary Kriging menghasilkan interpolasi yang sangat baik, yaitu rata – rata kuadrat error (MSE) yang dihasilkan sangat kecil atau sama dengan 0.
3. Kedua metode memberikan akurasi yang sama walau pun terdapat perbedaan yang sangat kecil.
4. Kedua metode dapat diterapkan untuk interpolasi data multivariate.

Saran

Untuk data sampel dengan ukuran kecil / sedikit sebaiknya digunakan interpolasi dengan metode Lagrange, juga sebaliknya untuk data sampel ukuran besar/banyak sebaiknya digunakan metode Ordinary Kriging.

Daftar Pustaka

1. Alex Lewis*, Multivariate Lagrange Interpolation.*
2. Armstrong, Margaret. 1998*. Basic Linear Geostatistics,* Springer: Germany.
3. Geoff Bohling, 2005*. Introduction to GeostatsisticsAndVariogram Analysis.*
4. *https://id.wikipedia.org/wiki/Kriging*
5. Isobel Clark, Elsevier Science, 1979*, Practical Geostatistics*
6. Kamron. Saniee, 2007.*A Simple Expression for Multivariate Lagrange Interpolation.*
7. Mitas, L., Mitasova, H., 1999*; Spatial Interpolation. In: P.Longley, et al (Eds.).*
8. P.J. Olver. 2006*.On multivariate interpolation. Studies in Applied Mathematics.*
9. TănaseDinu, SeriaMatematică - Informatică – Fizică, Vol. LIX No. 1/2007, *Interpolation of the Functions with Two Variable Values with Simple Nodes.*