Б1.Метод кригинга

Геостатистика, которая является разделом прикладной статистики, может характеризовать регулярную составляющую изменения природных объектов, включая грунты. Ее можно рассматривать как инструмент для изучения и прогнозирования структуры пространственно привязанных переменных. Основной инструмент геостатистики известный как анализ вариограмм, используется для определения и описания степени пространственной изменчивости региональных переменных. Кригинг, как геостатистический метод, основан на пространственной автокорреляции данных, которая определяет статистическую взаимосвязь между значениями, где доступны выборочные наблюдения. Когда связь установлена, она используется для прогнозирования значений переменных в местах без выборки.

Реализация метода кригинга требует большого количества вычислительного времени. Эффективность интерполяции кригинга зависит от выбора функций вариации в теоретической модели. Однако вариограмма, подходящая для геологической модели, заранее неизвестна, поэтому человеческий фактор в методе кригинга является значительным, поскольку пользователь определяет, является ли функция вариограммы разумной или нет.

Метод кригинга основан на использовании регионализированной переменной, т.е. переменной, которая изменяется от места к месту с некоторой видимой непрерывностью, поэтому не может моделироваться только одним математическим уравнением и как таковой, аналогичен по своей природе теории случайного поля. Регионализированная переменная обладает свойствами, которые частично случайны и частично пространственны, и имеет непрерывность от точки к точке, однако изменения настолько сложны, что их нельзя описать с помощью поддающейся обработке детерминированной функции.

В наиболее простом и самом распространенном варианте линейного кригинга оценка в произвольной точке p_0 рассматриваемой области ищется в виде линейной комбинации значений в опорных точках:

$$Z = a_1 z_1 + a_2 z_2 + \ldots + a_n z_n . \tag{1}$$

Неизвестные коэффициенты a_i определяются из условий несмещенности оценки и минимизации ее дисперсии, что приводит к системе линейных уравнений кригинга. При условии, что процесс Z имеет стационарное математическое ожидание, система уравнений Кригинга выглядит следующим образом

$$\sum_{j=1}^{n} a_j \gamma_{ij} + \mu = \gamma_i ,\; i = 1,\ldots,n \;;\; \sum_{j=1}^{n} a_j = 1 \;; \tag{2}$$

где μ - множитель Лагранжа, возникающий из-за условия несмещенности оценки, $\gamma_{ij} = \gamma(\boldsymbol{p}_i - \boldsymbol{p}_j)$, $\gamma_i = \gamma(\boldsymbol{p}_i - \boldsymbol{p}_0)$. Процедура оценивания в этом случае называется простым (обычным) кригингом (ordinary linear kriging). Так как минимизируемая дисперсия оценки может быть записана также и через ковариационную функцию, в уравнениях кригинга может использоваться функция *c(h)*. Если же процесс Z не является стационарным (в данных присутствует значимый тренд), используют процедуру универсального (universal) линейного кригинга. В этом случае считается, что случайный процесс Z может быть представлен в виде

$$Z(\boldsymbol{p}) = \sum_{i=1}^{k} b_i \varphi_i(\boldsymbol{p}) + \varepsilon(\boldsymbol{p}), \text{ где } \varphi_1(\boldsymbol{p}), \ldots, \varphi_k(\boldsymbol{p}) \quad (3)$$

где $\varphi_1(\boldsymbol{p}), \ldots, \varphi_k(\boldsymbol{p})$ - набор линейно-независимых детерминированных базисных функций, а $\varepsilon(\boldsymbol{p})$ - стационарный случайный процесс с нулевым математическим ожиданием. Детерминированная составляющая называется трендом, а в качестве функций $\phi_i(\boldsymbol{p})$ обычно используют полиномы, то есть тренд представляет собой полином, как правило, не выше второй степени. Минимизация дисперсии оценки и учет несмещенности приводят к системе уравнений универсального кригинга для коэффициентов a_i:

$$\begin{aligned} &\sum_{j=1}^{n} a_j \gamma_{ij} + \sum_{j=1}^{k} \mu_j \varphi_j(\boldsymbol{p}_i) = \gamma_i, i = 1, \ldots, n; \\ &\sum_{j=1}^{n} a_j \varphi_i(\boldsymbol{p}_j) = \varphi_i(\boldsymbol{p}_0), i = 1, \ldots, k; \end{aligned} \quad (4)$$

где μ и k ,..., 1 – множители Лагранжа.

Простой кригинг является частным случаем универсального при k =1 и ϕ_1 ($\boldsymbol{p}$) = 1. Метод кригинга, кроме вычисления самого показателя, позволяет найти дисперсию в точке p_0 по формуле

$$\sigma^2(\boldsymbol{p}_0) = \sum_{i=1}^{n} a_i \gamma_i + \sum_{i=1}^{k} \mu_i \varphi_i(\boldsymbol{p}_0), \quad (5)$$

Кригинг позволяет учесть вес каждой точки экспериментальных данных в процессе вычисления значений интерполяционной функции в точках регулярной сетки. Это один из наиболее гибких и часто используемых методов, однако на множествах большого размера он работает медленно.

Основная проблема кригинга – необходимость подбора параметров вариограммы, от которых зависит результат интерполяции. На первом этапе вычислений по исходным данным подбирается выборочная (экспериментальная) вариограмма $\gamma^*(h)$ и строится график функции, где каждому интервалу значений *h* (разность расстояний между парами точек;

имеет размерность в единицах карты – градусах) соответствует вариация (квадрат разности значений величин в этих точках); h откладывается по оси x или y и предполагается, что поле изотропно.

Выборочную вариограмму нельзя напрямую использовать в уравнениях кригинга, её необходимо приблизить некоторой модельной функцией вариограммы *γ(h)* (см., **вариограммы**), которая используется на втором этапе вычислений. Наилучшим вариантом подобранной модельной функции можно считать тот, который дает наименьшую дисперсию отклонений эмпирических значений от теоретических.

Оценки кригинга рассчитываются как взвешенные суммы соседних образцов. Эти веса зависят от представленной структуры корреляции. Например, если данные кажутся очень непрерывными в пространстве, те точки, которые находятся ближе к расчетным точкам, получают более высокие веса, чем те, которые находятся дальше. Критерием выбора этих весов является минимизация дисперсии оценки.