Лекция 15

Случайные процессы

Параметрические модели временных рядов. Сглаживание и фильтрация. Моделирование искусственных гидрологических рядов

(Ахметов С.К.)

Параметрические модели временных рядов

- ***Непараметрические методы*** - это методы описания СП с помощью корреляционных функций *(это то, что делалось до сих пор)*
- ***Параметрические методы*** – описание СП с помощью моделей авторегрессии и их комбинаций

Модель авторегрессии

- В этой модели текущие значения СП выражаются в виде линейной комбинации предыдущих его значений и белого шума:

$$X'(t) = \varphi_1 X'(t-1) + \varphi_2 X'(t-2) + \ldots\ldots \varphi_p X'(t-p) + a(t)$$

$X'(t)$ – центрированный СП; $X'(t) = X(t) - m_x$

$a(t)$ – белый шум с нулевым математическим ожиданием и среднеквадратическим отклонение σ_a

$\varphi_1, \varphi_2 \ldots\ldots \varphi_p$ – коэффициенты модели (константы)

m_x – математическое ожидание

Параметрические модели временных рядов (2)

Модель содержит $p + 2$ параметров: $\varphi_{1,} \varphi_{2}\ldots\ldots \varphi_{p,} m_{x,} \sigma_{a}$

Эта модель авторегрессии p-го порядка. Она обозначается как **АР(p)**

При $p=1$

$$X'(t) = \varphi_1 X(t-1) + a(t)$$

Это модель называется моделью авторегрессии первого порядка *АР(1)* или модель Марковского процесса. Для этой модели коэффициент φ_1 и ординаты автокорреляционной функции связаны соотношением

$$r_k = \varphi_1 r_{k-1} \text{ при } k > 0$$

Так как $r_0 = 1$, то

$$r_k = \varphi_1^{\,k}$$

Таким образом, для *АР(1)* автокорреляционная функция полностью определяется своей первой ординатой. При этом $\varphi_1 = r_1$

Модель скользящего среднего

- Модель скользящего среднего м.б. получена из общей линейной модели (ОЛМ) при предположении, что ОЛМ содержит конечное число членов. При этом текущие значения ***СП X(t)*** выражаются через предыдущие значения белого шума $a_{t-1}, a_{t-2} \ldots a_{t-q}$

$$X'(t) = a(t) - \eta_1 a(t-1) - \eta_2 a(t-2) - \ldots - \eta_q a(t-q)$$

- Модель содержит $q + 2$ параметров: $\eta_1, \eta_2 \ldots - \eta_q, m_x, \sigma_a$

$X'(t)$ – центрированный случайный процесс, $X'(t) = X(t) - m_x$

$a(t)$ - белый шум с нулевым математическим ожиданием и среднеквадратическим отклонение σ_a

$\eta_1, \eta_2 \ldots - \eta_q$ - коэффициенты модели (константы)

Это выражение называется моделью скользящего среднего q - го порядка и обозначается ***СС(q)***

Смешанные модели

Иногда целесообразно объединять модели АР и СС. В этом случае получается смешанная модель ***ССАР (p,q)***, где ***p*** – порядок авторегрессии, ***q*** - порядок скользящего среднего. Выражение ***ССАР*** имеет вид

$$X'(t) = \varphi_1 X'(t-1) + \varphi_2 X'(t-2) + \ldots\ldots \varphi_p X'(t-p) + a(t) + a(t) - \eta_1\, a(t-1) - \eta_2\, a(t-2) - \ldots - \eta_q\, a(t-q)$$

Такая модель может быть полезна, например, когда наблюдаемый временной ряд является суммой двух или более независимых составляющих, каждая из которых описывается либо моделью АР, либо моделью СС, но которые непосредственно не измеряются. При ***p = 1*** и ***q=1***, модель имеет вид

$$X'(t) = \varphi_1 X'(t-1) + a(t) - \eta_1\, a(t-1)$$

Модель авторегрессии и проинтегрированного скользящего среднего (АРПСС)

- Модели АР, СС и АРСС относятся к классу ***стационарных*** моделей, которые описывают процессы с постоянными МО и дисперсиями
- Если ряды нестационарны, то от нестационарности можно избавиться, заменяя исходный ряд на ряд разностей:

$$Y(t) = X(t) - X(t-1)$$

- Если от нестационарности не удается избавиться, то можно взять разность повторно
- После проведенных преобразований к исходному ряду можно применить модель АРСС
- Такую модель называют моделью АРПСС (p,d,q). В этой модели:

p - порядок авторегрессии
d - порядок разности
q - порядок скользящего среднего

Сглаживание и фильтрация

- Если временной ряд содержит некоторые частоты и периоды, которые в данный момент не представляют интереса, то амплитуда этих волн может быть уменьшена с помощью статистической фильтрации. При этом изменяется спектр исходного ряда.

- Одной из форм статистической фильтрации может быть сглаживание, в которой спектральные компоненты с высокой частотой уменьшены. Такой фильтр называется ***низкочастотным.***

- Простейшим статистическим фильтром является скользящая средняя с равными весами. Скользящее среднее рассчитывается путем суммирования ***n*** последовательных величин временного ряда и делением полученной суммы на ***n*.**

Оценка характеристик СП по эмпирическим данным

Определение характеристик СП по множеству реализаций

Если имеется ***k*** реализаций СП, тогда ***МО*** выражается формулой

$$m_x^*(t_i) = \frac{1}{k}\sum_{j=1}^{k} x_j(t_i)$$

Таким же образом оценивается дисперсия. То есть нужно рассчитать оценку дисперсии для каждого сечения

$$D_x^*(t_i) = \frac{1}{k-1}\sum_{j=1}^{j=k}[x_j(t_i) - m_x^*(t_i)]^2$$

Оценка ***СКО*** ***СП*** равна

$$S_x(t_i) = \sqrt{D_x^*(t_i)}$$

Оценка корреляционной функции СП определяется выражением

$$K_x^*(t_i, t_q) = \frac{1}{k-1}\sum_{j=1}^{j=k}[x_j(t_i) - m_x^*(t_i)][x_j(t_q) - m_x^*(t_q)]$$

Оценка нормированной корреляционной функции СП определяется выражением

$$r_x^*(t_i, t_q) = \frac{K_x^*(t_i, t_q)}{S_x(t_i)S_x(t_q)}$$

Оценка характеристик СП по эмпирическим данным

- Для стационарного СП ***математическое ожидание, дисперсия*** и ***СКО*** являются константами и их можно оценить по любому сечению.
- Корреляционная функция стационарного процесса не зависит от моментов времени рассматриваемого процесса и зависит только от расстояния между сечениями τ. Поэтому ее оценивают по формуле

$$k_x^*(\tau) = \frac{1}{k-1}\sum_{j=1}^{j=k}[x_j(t) - m_x^*][x_j(t+\tau) - m_x^*]$$

- Оценка нормированной корреляционной функции стационарного процесса имеет вид

$$r_x^*(\tau) = \frac{K_x^*(\tau)}{S_x^2}$$

Эту функцию называют автокорреляционной функцией

Определение характеристик стационарного эргодического процесса по одной реализации

- Если для стационарного СП принимается гипотеза об эгродичности, то ***МО, дисперсию*** и ***СКО*** можно оценить по одной (достаточно продолжительной) реализации. Тогда получим, что

$$m_x^*(t) = m_x^* = \frac{1}{n}\sum_{i=1}^{i=n} x_i \qquad D_x^*(t) = D_x^* = \frac{1}{n-1}\sum_{i=1}^{i=n}(x_i - m_x^*)^2$$

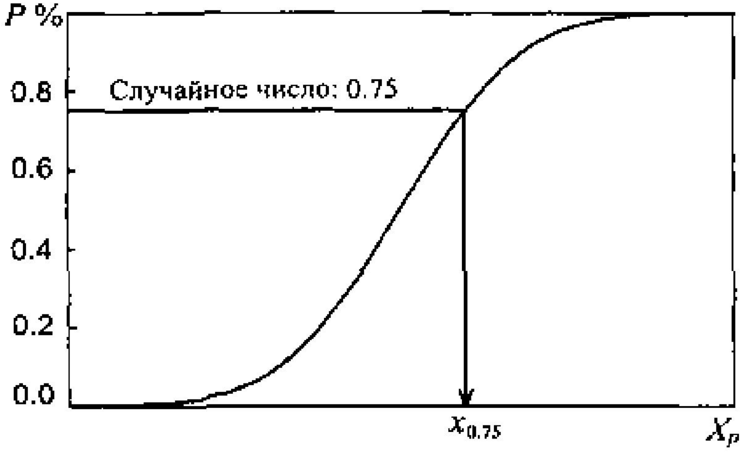
$$\sigma_x^*(t) = \sigma_x^* = \sqrt{\frac{1}{n-1}\sum_{i=1}^{i=n}(x_i - m_x^*)^2}$$

- В практике гидрологических расчетов часто используется первая ордината автокорреляционной функции, которая называется ***коэффициентом автокорреляции.*** Он определяется выражением

$$r_x^*(1) = \frac{\sum_{i=1}^{i=n-1}(x_i - m_x^*)(x_{i+1} - m_x^*)}{(n-2)D_x^*}$$

Моделирование искусственных гидрологических рядов

- В основе моделирования искусственных рядов лежит метод Монте – Карло. Это метод решения математических задач при помощи моделирования случайных чисел
- Процесс моделирования включает следующие шаги:
- Нужно получить последовательность случайных чисел, равномерно распределенных в интервале от 0 до 1
- Каждое значение случайного числа рассматривается как вероятность не превышения и по нему рассчитывается соответствующий квантиль заданного закона распределения. Можно это сделать по графику, можно по таблицам или по готовым компьютерным программа.



СПАСИБО ЗА ВНИМАНИЕ!