

Лекция 8: Метод группового учёта аргументов (МГУА)

- Метод наименьших квадратов
- Общая схема алгоритмов МГУА
- Алгоритм с ковариациями и квадратичными описаниями

Метод наименьших квадратов

Перед тем, как начинать рассмотрение МГУА, было бы полезно вспомнить или узнать впервые метод наименьших квадратов — наиболее распространенный метод подстройки линейно зависимых параметров.

Рассмотрим для примера МНК для трех аргументов:

Пусть функция $T=T(U, V, W)$ задана таблицей, то есть из опыта известны числа U_i, V_i, W_i, T_i ($i = 1, \dots, n$). Будем искать зависимость между этими данными в виде:

$$T(U, V, W) = aU + bV + cW \quad (1)$$

где a, b, c — неизвестные параметры.

Подберём значения этих параметров так, чтобы была наименьшей сумма квадратов отклонений опытных данных T_i и теоретических $T_i = aU_i + bV_i + cW_i$, то есть сумма:

$$\sigma = \sum_{i=1}^n (T_i - aU_i - bV_i - cW_i)^2 \rightarrow \min \quad (2)$$

Величина σ является функцией трех переменных a, b, c . Необходимым и достаточным условием существования минимума этой функции является равенство нулю частных производных функции σ по всем переменным, то есть:

$$\frac{\partial \sigma}{\partial a} = 0, \frac{\partial \sigma}{\partial b} = 0, \frac{\partial \sigma}{\partial c} = 0 \quad (3)$$

Так как:

$$\begin{aligned} \frac{\partial \sigma}{\partial a} &= -2 \sum_{i=1}^n (T_i - aU_i - bV_i - cW_i)U_i \\ \frac{\partial \sigma}{\partial b} &= -2 \sum_{i=1}^n (T_i - aU_i - bV_i - cW_i)V_i \\ \frac{\partial \sigma}{\partial c} &= -2 \sum_{i=1}^n (T_i - aU_i - bV_i - cW_i)W_i \end{aligned} \quad (4)$$

то система для нахождения a , b , c будет иметь вид:

$$\begin{aligned} a \sum_{i=1}^n U_i^2 + b \sum_{i=1}^n U_i V_i + c \sum_{i=1}^n U_i W_i &= \sum_{i=1}^n T_i U_i \\ a \sum_{i=1}^n U_i V_i + b \sum_{i=1}^n V_i^2 + c \sum_{i=1}^n V_i W_i &= \sum_{i=1}^n T_i V_i \\ a \sum_{i=1}^n U_i W_i + b \sum_{i=1}^n W_i V_i + c \sum_{i=1}^n W_i^2 &= \sum_{i=1}^n T_i W_i \end{aligned} \quad (5)$$

Данная система решается любым стандартным методом решения систем линейных уравнений (Гаусса, Жордана, Зейделя, Крамера).

Рассмотрим некоторые практические примеры нахождения приближающих функций:

$$y = ax^2 + bx + g$$

Задача подбора коэффициентов a, b, g сводится к решению общей задачи при $T = y, U = x^2, V = x, W = 1, a = a, b = b, g = c$

$$f(x, y) = a \sin(x) + b \cos(y) + \frac{g}{x}$$

Задача подбора коэффициентов a, b, g сводится к решению общей задачи при $T = f, U = \sin(x), V = \cos(y), W = 1/x, a = a, b = b, g = c$

Если мы распространим МНК на случай с m параметрами,

$$\sigma = \sum_{i=1}^n (T_i - \sum_{v=1}^m u_{iv} c_v)^2 \rightarrow \min \quad (6)$$

то путем рассуждений, аналогичных приведенным выше, получим следующую систему линейных уравнений:

$$\begin{cases} \overline{c_1 u_1 u_1} + \overline{c_2 u_1 u_2} + \dots + \overline{c_m u_1 u_m} = \overline{T u_1} \\ \overline{c_1 u_2 u_1} + \overline{c_2 u_2 u_2} + \dots + \overline{c_m u_2 u_m} = \overline{T u_2} \\ \dots \\ \overline{c_1 u_m u_1} + \overline{c_2 u_m u_2} + \dots + \overline{c_m u_m u_m} = \overline{T u_m} \end{cases} \quad (7)$$

где $\overline{T} = \{T_i\}_{i=1}^m$, $\overline{u}_v = \{u_{iv}\}_{i=1}^n$

Общая схема построения алгоритмов метода группового учета аргументов (МГУА).

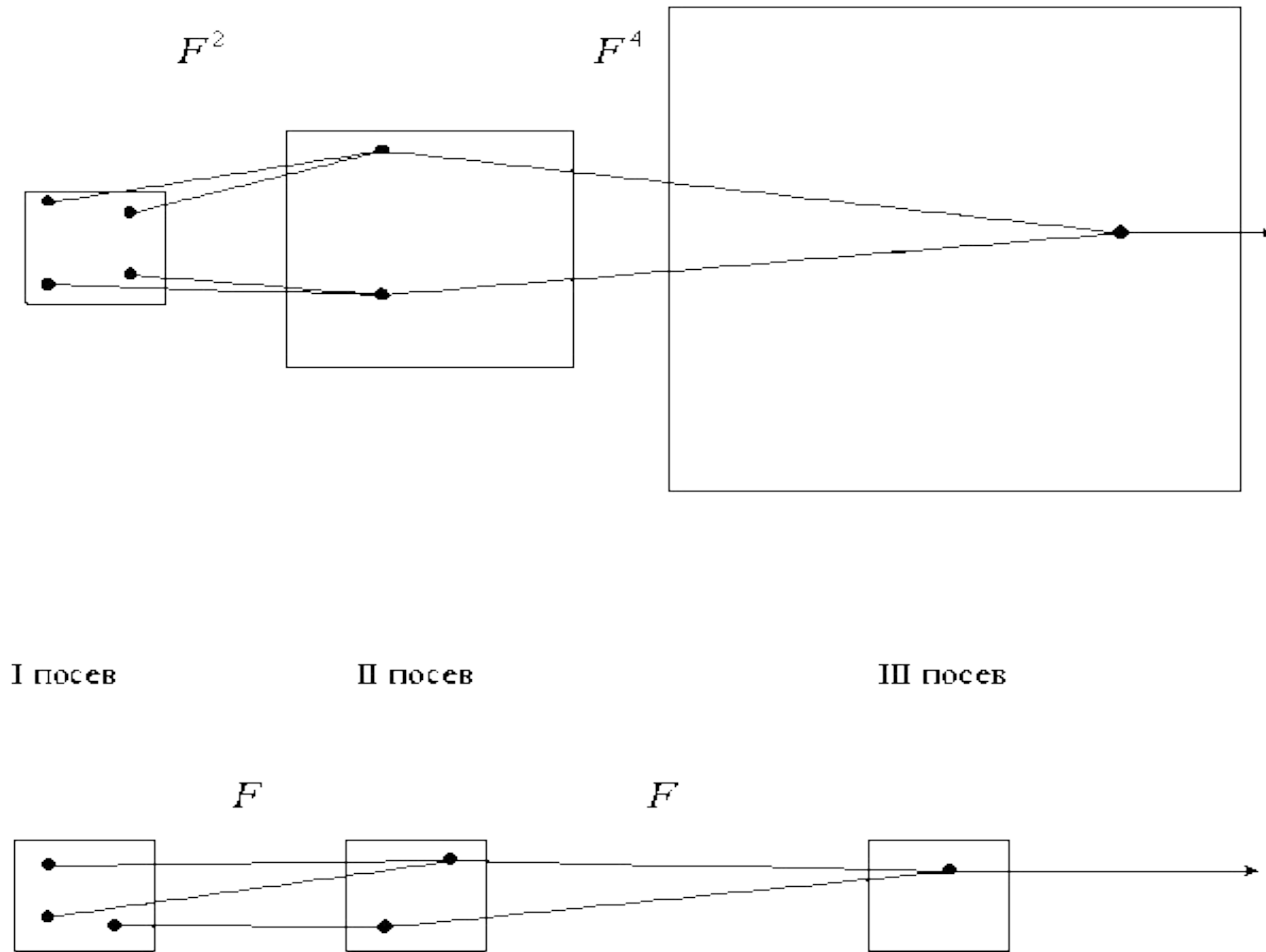


Рис. 1. Селекция самого черного тюльпана при расширяющемся опытном поле (эквивалент полного перебора), и при постоянном размере поля (эквивалент селекции при сохранении свободы выбора решений $F = \text{const}$).

Алгоритмы МГУА воспроизводят схему массовой селекции [Васильев В. И., Коноваленко В. В., Горелов Ю. И. Имитационное управление неопределенными объектами. // К.: “Наукова думка”, 1989—216с.], показанной на Рис. 1.

В них есть генераторы усложняющихся из ряда в ряд комбинаций и пороговые самоотборы лучших из них. Так называемое “полное” описание объекта

$$\sigma = f(x_1, x_2, x_3, \dots, x_m)$$

где f — некоторая элементарная функция, например степенной полином, заменяется несколькими рядами “частных” описаний:

$$\text{1-ряд селекции: } y_1 = f(x_1x_2), y_2 = f(x_1x_3), \dots, y_s = f(x_{m-1}x_m)$$

$$\text{2-ряд селекции: } z_1 = f(y_1y_2), z_2 = f(y_1y_2), \dots, z_p = f(y_{s-1}y_s)$$

Входные аргументы и промежуточные переменные сопрягаются попарно, и сложность комбинаций на каждом ряду обработки информации возрастает (как при массовой селекции), пока не будет получена единственная модель оптимальной сложности.

Каждое частное описание является функцией только двух аргументов. Поэтому его коэффициенты легко определить по данным обучающей последовательности при малом числе узлов интерполяции.

Исключая промежуточные переменные (если это удастся), можно получить "аналог" полного описания. Математика не запрещает обе эти операции. Например, по десяти узлам интерполяции можно получить в результате оценки коэффициентов полинома сотой степени и т. д.

Из ряда в ряд селекции пропускается только некоторое количество самых регулярных переменных. Степень регулярности оценивается по величине среднеквадратичной ошибки (средней для всех выбираемых в каждом поколении переменных или для одной самой точной переменной) на отдельной проверочной последовательности данных.

Иногда в качестве показателя регулярности используется коэффициент корреляции.

Ряды селекции наращиваются до тех пор, пока регулярность повышается. Как только достигнут минимум ошибки, селекцию следует остановить. Практически рекомендуется остановить селекцию даже несколько раньше достижения полного минимума, как только ошибка начинает падать слишком медленно. Это приводит к более простым и более достоверным уравнениям.

Алгоритм с ковариациями и с квадратичными описаниями.

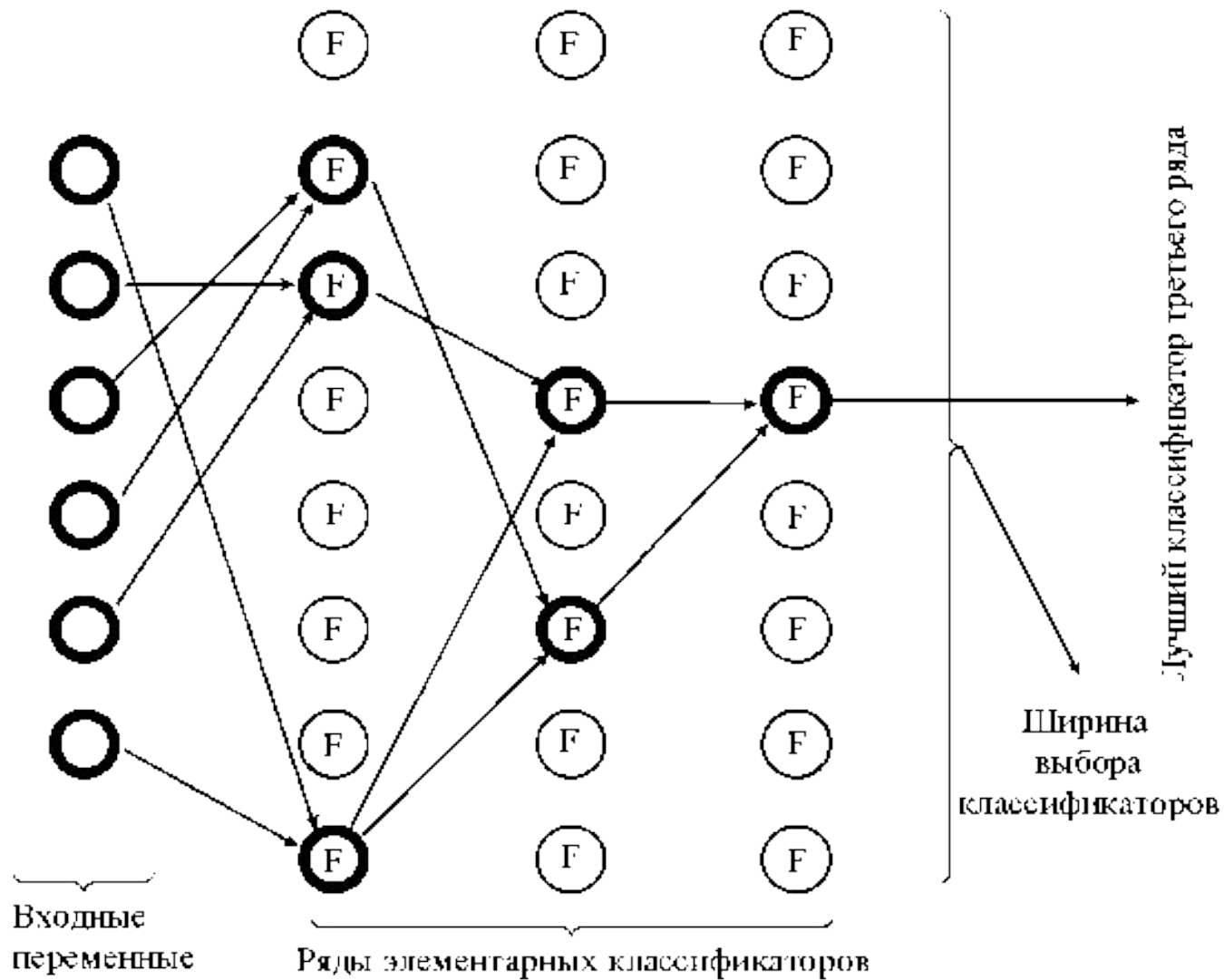


Рис. 2

В этом алгоритме [Васильев В. И., Коноваленко В. В., Горелов Ю. И. Имитационное управление неопределенными объектами. // К.: "Наукова думка", 1989—216с. , Половинкин А. И. Основы инженерного творчества. // М.: "Машиностроение", 1988—368с] используются частные описания, представленные в следующих формулах:

$$y_i = a_0 + a_1x_i + a_2x_j + a_3x_ix_j \quad (8)$$

$$y_k = a_0 + a_1x_i + a_2x_j + a_3x_ix_j + a_4x_i^2 + a_5x_j^2$$

Сложность модели увеличивается от ряда к ряду селекции как по числу учитываемых аргументов, так и по степени.

Степень полного описания быстро растет. На первом ряду — квадратичные описания, на втором — четвертой степени, на третьем — восьмой и т. д.

В связи с этим минимум критерия селекции находится быстро, но не совсем точно.

Кроме того, имеется опасность потери существенного аргумента, особенно на первых рядах селекции (в случае отсутствия протекции).

Специальные теоремы теории МГУА определяют условия, при которых результат селекции не отличается от результата полного перебора моделей.

Для того чтобы степень полного уравнения повышалась с каждым рядом селекции на единицу, достаточно рассматривать все аргументы и их ковариации как обобщенные аргументы и пользоваться составленными для них линейными описаниями.