Метод на най-малките квадрати

Разглеждаме ***M*** – множество от функции $f(x)$, зададени таблично в N - точки (не задължително различни) и $P \equiv \prod_n(x)$ – полиномите от n -та степен на променливата x. Ще считаме, че ***n* << *N*.** Като мярка за близост между функцията $f(x)$ от множеството M и $P(x) \in \prod_n(x)$ ще използваме стойностите на функцията:

$$\Phi(a_0,...,a_n) = \sum_{i=1}^{N}[y_i - (a_n x_i^n + a_{n-1}x_i^{n-1} + ... + a_1 x_i + a_0)]^2,$$

където $a_0,...,a_n$ са коефициентите на полинома $P(x)$, а $y_i = f(x_i)$. Полиномът P^*, чиито коефициенти $a_0^*,...,a_n^*$ минимизират функцията $\Phi(a_0,...,a_n)$, се нарича полином на най-добро приближение по метода на най-малките квадрати (**МНМК**) и може да бъде използван като приближение на $f(x)$ (особено когато N е много по-голямо от n). Коефициентите $a_0^*,...,a_n^*$ са решение на следната линейна алгебрична система (която е със симетрична матрица и за нейното решаване може да се използва методът на квадратния корен – виж параграф 1.5):

$$\left|\begin{array}{l}
Na_0 + \left(\sum_{i=1}^{N} x_i\right)a_1 + \left(\sum_{i=1}^{N} x_i^2\right)a_2 + ... + \left(\sum_{i=1}^{N} x_i^n\right)a_n = \sum_{i=1}^{N} y_i \\
\left(\sum_{i=1}^{N} x_i\right)a_0 + \left(\sum_{i=1}^{N} x_i^2\right)a_1 + \left(\sum_{i=1}^{N} x_i^3\right)a_2 + ... + \left(\sum_{i=1}^{N} x_i^{n+1}\right)a_n = \sum_{i=1}^{N} x_i y_i \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \\
\left(\sum_{i=1}^{N} x_i^n\right)a_0 + \left(\sum_{i=1}^{N} x_i^{n+1}\right)a_1 + \left(\sum_{i=1}^{N} x_i^{n+2}\right)a_2 + ... + \left(\sum_{i=1}^{N} x_i^{2n}\right)a_n = \sum_{i=1}^{N} x_i^n y_i
\end{array}\right.$$

По аналогия с метода на най-малките квадрати за приближаване на таблично зададени функции се въвежда понятието **“решение по метода на най-**

малките квадрати" за преопределени системи линейни алгебрични уравнения (броят на уравненията m е по-голям от броя на неизвестните n):

Ако преопределената система има вида: $Ax = b$:

$$\left| \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n \end{array} \right. \quad \text{при} \quad m > n.$$

Под решение по МНМК на тази система ще разбираме n – торката $(x_1^*, \ldots, x_n^*)$, която минимизира израза:

$$\Phi(x_1, \ldots, x_n) = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n - b_1)^2 + (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n - b_2)^2 + \\ \ldots + (a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n - b_m)^2$$

Точката, която минимизира тази функция е решение на квадратната система, която се получава като умножим отляво изходната система с A^T :$\boldsymbol{A^T A x} = \boldsymbol{A^T b}$, което се нарича още симетризиране на системата.

Пример 1. Да се намерят P_1^* и P_2^* по МНМК за таблично зададена функция $f(x)$:

x_i	0	1	2	3	4
$y_i = f(x_i)$	1	2	1	0	4

Решение:

За намиране на P_1^* построяваме таблицата от стойностите на $x_i, y_i, x_i^2, y_i x_i$ и намираме необходимите суми:

i	x_i	y_i	x_i^2	$x_i y_i$
1	0	1	0	0
2	1	2	1	2
3	2	1	4	2
4	3	0	9	0
5	4	4	16	16
Σ	**10**	**8**	**30**	**20**

Тогава, ако $P_1^* = a_1^* x + a_0^*$, коефициентите a_0^* и a_1^* са решение на системата:

$$\begin{cases} 5a_0 + 10a_1 = 8 \\ 10a_0 + 30a_1 = 20 \end{cases},$$

чиито решения са $a_0^* = \frac{4}{5}$ и $a_1^* = \frac{2}{5} \Rightarrow P_1^* = \frac{2}{5}x + \frac{4}{5}$.

За намиране на полинома от втора степен P_2^* допълваме горната таблица с още три колони: $x_i^3, x_i^4, y_i x_i^2$:

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$y_i x_i^2$
1	0	1	0	0	0	0	0
2	1	2	1	1	1	2	2
3	2	1	4	8	16	2	4
4	3	0	9	27	81	0	0
5	4	4	16	64	256	16	64
Σ	**10**	**8**	**30**	**100**	**354**	**20**	**70**

И системата изглежда така:

$$\begin{cases} 5a_0 + 10a_1 + 30a_2 = 8 \\ 10a_0 + 30a_1 + 100a_2 = 20 \\ 30a_0 + 100a_1 + 354a_2 = 70 \end{cases}$$

Решенията на системата (с точност до пет знака) са: $a_0^* = 1{,}65714$; $a_1^* =$-1,31429; $a_2^* = 0{,}42857$ и $P_2^*(x) = 0{,}42857x^2 - 1{,}31429x + 1{,}65714$.

Пример 2. По МНМК да се реши преопределената система:

$$\begin{cases} x + y = 2 \\ x - y = 0 \\ 3x + y = 3 \end{cases}.$$

Решение:

Матрицата $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 1 \end{pmatrix}$; $A^T = \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & 1 \end{pmatrix}$ и векторът $b = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$. Тогава

$$A^T A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 11 & 3 \\ 3 & 3 \end{pmatrix} \; ; \; A^T b = \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \end{pmatrix}.$$

И получаваме симетризираната система: $\begin{cases} 11x + 3y = 11 \\ 3x + 3y = 5 \end{cases}$.

След изваждане на второто уравнение от първото имаме $8x = 6 \;\Rightarrow\; x = \frac{3}{4}$ и след заместване $3y = 5 - \frac{9}{4} \;\Rightarrow\; y = \frac{11}{12}$, т.е. решението на преопределената система са $x^* = \frac{3}{4}$ и $y^* = \frac{11}{12}$.

Автор: Дойчин Бояджиев, dtb@uni-plovdiv.bg