Переопределенные системы линейных уравнений

Скалько Юрий Иванович
Цыбулин Иван
Шевченко Александр

Переопределенные СЛАУ

Рассмотрим СЛАУ
$$\mathbf{Ax} = \mathbf{b},$$
но в случае, когда уравнений больше, чем неизвестных. В этом случае система называется переопределенной. Почти всегда переопределенная система не имеет точного решения.

Решение переопределенной СЛАУ

Поскольку точного решения такая система почти никогда не имеет, необходимо ввести другое понятие решения переопределенной СЛАУ. Введем, как и раньше, невязку $\mathbf{r} = \mathbf{A}\mathbf{x} - \mathbf{b}$. Для точного решения требовалось $\mathbf{r} = 0$. Назовем решением переопределенной системы такое значение $\mathbf{x}$ для которого невязка минимальна

$$\|\mathbf{r}\|_E^2 \to \min_{\mathbf{x}}$$

Такое решение называется решением по методу наименьших квадратов

Метод наименьших квадратов

Запишем квадрат нормы невязки

$$\|\mathbf{r}\|_E^2 = \mathbf{r}^T\mathbf{r} = (\mathbf{A}\mathbf{x} - \mathbf{b})^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{A}^T\mathbf{b} + \mathbf{b}^T\mathbf{b}$$

Минимум достигается в точке, где градиент этой функции обращается в ноль, то есть

$$\nabla\|\mathbf{r}\|_E^2 = 2(\mathbf{A}^T\mathbf{A}\mathbf{x} - \mathbf{A}^T\mathbf{b}) = 0$$

Рассмотрим систему

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

Это система с квадратной матрицей, которая не вырождена, если ранг матрицы $\mathbf{A}$ равен числу неизвестных. Решение этой системы будет решением по методу наименьших квадратов.

Метод наименьших квадратов

Метод широко применяется при анализе экспериментальных данных для определения параметров зависимостей, проведения наилучших прямых и т.п. Пусть имеется набор данных $(x_i, y_i), i = \overline{1, n}$ и требуется по ним построить наилучшую прямую. Запишем уравнение прямой

$$y = \alpha x + \beta$$

Здесь α, β - неизвестные, причем на эти неизвестные имеется n уравнений

$$y_i = \alpha x_i + \beta, i = \overline{1, n}$$

$$y_i = \alpha x_i + \beta, i = \overline{1, n}$$

Запишем эту систему в матричной форме

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & 1 \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Решая ее по методу наименьших квадратов, необходимо решить систему

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & 1 \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Метод наименьших квадратов

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & 1 \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Перемножая матрицы, получим

$$\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

$$\alpha = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - \left(\sum x_i\right)^2}$$

$$\beta = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - \left(\sum x_i\right)^2}$$

Более сложные зависимости

Методом наименьших квадратов можно искать наилучшие коэффициенты и в нелинейных зависимостях. Важно, чтобы искомые коэффициенты входили в зависимость линейно. Например, рассмотрим зависимость

$$y = \alpha + \beta x + \gamma x^3$$

Сама зависимость нелинейна, но коэффициенты α, β, γ входят в нее линейно.

Более сложные зависимости

Аналогично

$$y_i = \alpha + \beta x_i + \gamma x_i^3, i = \overline{1, n}$$

Запишем эту систему в матричной форме

$$\begin{pmatrix} 1 & x_1 & x_1^3 \\ 1 & x_2 & x_2^3 \\ 1 & \vdots & \vdots \\ 1 & x_n & x_n^3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Решая ее по методу наименьших квадратов, необходимо решить систему

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^3 & x_2^3 & \dots & x_n^3 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^3 \\ 1 & x_2 & x_2^3 \\ 1 & \vdots & \vdots \\ 1 & x_n & x_n^3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^3 & x_2^3 & \dots & x_n^3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Более сложные зависимости

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^3 & x_2^3 & \dots & x_n^3 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^3 \\ 1 & x_2 & x_2^3 \\ 1 & \vdots & \vdots \\ 1 & x_n & x_n^3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^3 & x_2^3 & \dots & x_n^3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Умножая матрицы

$$\begin{pmatrix} n & \sum x_i & \sum x_i^3 \\ \sum x_i & \sum x_i^2 & \sum x_i^4 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^6 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^3 y_i \end{pmatrix}$$

Решение последней системы дает наилучшие значения α, β, γ.

Нелинейные зависимости от искомых коэффициентов

В некоторых случаях неизвестные коэффициенты входят в зависимости нелинейно. Иногда можно изменить уравнение так, чтобы коэффициенты входили линейно

$$\sin(\omega x_i + \varphi) = y_i \to \omega x_i + \varphi = \arcsin y_i$$
$$\rho_i = cT_i^\gamma \to \ln \rho_i = \ln c + \gamma \ln T_i$$

Во втором случае неопределенными коэффициентам будут $\xi \equiv \ln c$ и γ. После нахождения ξ можно легко найти $c = e^\xi$

Взвешенный метод наименьших квадратов

Допустим, по некоторым причинам, некоторым уравнениям в системе приписан больший вес в суммарной невязке. Это означает, что одни уравнения должны выполняться точнее других, или наоборот, некоторые уравнения вообще не должны влиять на результат.
Пусть дана диагональная матрица $\mathbf{W} = \mathbf{W}^T > 0$. Вместо минимизации $\|\mathbf{r}\|_E^2 = \mathbf{r}^T\mathbf{r}$ будем минимизировать $\|\mathbf{r}\|_W^2 \equiv \mathbf{r}^T\mathbf{W}\mathbf{r}$. Чем больше значение W_{ii} тем точнее должно выполняться уравнение i (его невязка учитывается с большим весом)

$$\mathbf{r}^T\mathbf{W}\mathbf{r} = (\mathbf{x}^T\mathbf{A}^T - \mathbf{b}^T)\mathbf{W}(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{x}^T\mathbf{A}^T\mathbf{W}\mathbf{A}\mathbf{x} - 2\mathbf{b}^T\mathbf{W}\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{b}$$

Аналогично, получаем обычную СЛАУ

$$\mathbf{A}^T\mathbf{W}\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{W}\mathbf{b}$$

Взвешенный метод наименьших квадратов

Пусть имеются те же самые данные (x_i, y_i), но еще дополнительно известна погрешность δy_i (например, погрешность измерений). Имеет смысл вместо задачи

$$\sum_i (y_i - \alpha x_i - \beta)^2 \to \min_{\alpha,\beta}$$

решать задачу

$$\sum_i \left(\frac{y_i - \alpha x_i - \beta}{\delta y_i}\right)^2 \to \min_{\alpha,\beta}$$

Тогда матрица $\mathbf{W} = diag\left(\frac{1}{\delta y_1^2}, \ldots, \frac{1}{\delta y_n^2}\right)$. Чем больше ошибка измерения, тем меньше «вес» уравнения.

Спасибо за внимание!

tsybulinhome@gmail.com