УДК 681.3:518.1

Применение метода линеаризации для генерации регрессионных моделей в материаловедении

Мельниченко В.В., Стефанович В.А.
Белорусский национальный технический университет

Целью настоящей работы является анализ технологий разработки регрессионных моделей, с целью выработки рекомендаций получения регрессионных моделей для сильно анизотропных сред.

Для решения такой задачи надо сделать первый шаг (выбор). Исследователь находится в точке А и в сторону точки В делает пробный шаг, затем второй и так далее. Обозначим точку А(x,y), здесь X,Y оси системы координат. В простейшем случае прямоугольная. Тогда траектория движения есть набор точек В(x,y). Рассмотрим эту траекторию движения. Если обозначить номер точки буквой i, получим два массива координат x_i, y_i. Расстояние между соседними точками установим одинаковым, и они приблизительно ложатся на одну прямую линию, которую исследователь проводит из последней точки в первую [1]. Регрессионная зависимость установлена как линейная.

$$\varphi(x) = k\phi(x) + d \qquad (1)$$

В общем случае эта зависимость может быть и нелинейной. Если точки не ложатся на линию, исследователь пробует предположить другие зависимости, рассматриваемые как простейшие [2]

На плоскости в системе координат XY эта функция отображается в виде кривой линии. Предположим, что обе функции из уравнения (1) строго монотонные, тогда возможно введение преобразование переменных

$$X1 = \phi(x), Y1 = \varphi(x) \qquad (2)$$

В новой системе координат точки будут располагаться на прямой линии. Это дает возможность визуальной оценки типа зависимости как линейной, если представить результаты в преобразованной системе координат.

Ряд простейших эмпирических зависимостей линеаризуется логарифмированием по одной или обеим координатным осям [1].

Этот метод легко распространяется и на многомерное пространство при исследовании многокомпонентного материала (сплава).

В случае, когда логарифмические преобразования трудоемки предложен метод наименьших квадратов отклонений предсказываемой регрессионной моделью от опытной точки. Накопленная ошибка является мерой оценки для модели в целом. Та модель, которая дает наименьшую сумму квадратов рассматривается как наилучшая. Чтобы доказать ее применимость выбранной регрессионной модели вычисляется критерий Фишера[3], который является дробью. В числителе дроби стоит сумма квадратов отклонений, а в знаменателе оценка воспроизводимости точек эксперимента. Следуя закону нормального распределения ошибок составленные таблицы являются критерием достоверности предлагаемой регрессионной модели.

Вывод гибридных корреляционных моделей, которые состоят из двух базовых регрессионных моделей и оценка адекватности примененной модели в пакете Mathcad [3] реализуется применением специальных процедур linfit(x,y,F(x)) и genfit(x,y,F(…),v). Но и имея такую программную поддержку подбор необходимой, корректной регрессионной модели является искусством экспериментатора, поскольку требуется введение коррелирующих коэффициентов регрессионные модели.

Встает единственная задача обучение вычислительного процесса методу аппроксимации имеющихся данных. В задачах с разрывами первого рода (например, вычисление формы, скорости роста фаз в многокомпонентной среде) необходимым условием с нашей точки зрения является применение метода нейронных сетей. В этом случае мелкие шаги, как бы нащупывают границы разрывов с обеих сторон и компромиссно создают границу. Наиболее простой программной оболочкой, в которой можно реализовать данную концепцию, можно рассматривать программу Statistica [4].

Литература

1. Демидович, Б.П. Численные методы анализа. // Б.П. Демидович, И.А. Марон, Э.З. Шувалова. – М.: Наука, 1963.- 400 с.
2. Тарасевич, Ю.Ю. Численные методы на Mathcad. //Астраханский университет, Астрахань, 2000. – 68 с.
3. Тейлор, Дж. Введение в теорию ошибок. // Дж. Тейлор. М.: Мир, 1985.- 288 с.
4. Statistica. Официальное руководство. Том3. – 2007.