Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

Отчет по лабораторной работе № 1
по эконометрике
"Постороение регрессионной модели, линейной по параметрам"

Выполнила студентка:
Заболотских Екатерина
группа: 3630102/70301

Проверил:
к.ф.-м.н., доцент
Иванков Алексей Александрович

Санкт-Петербург
2021 г.

Содержание

1 Постановка задачи

Имеется реализация случайного процесса. Задача состоит в построении регрессионной модели с разными апроксимирующими полиномами. Нужно найти оценки параметров β с помощью минимизации квадратичной ошибки МНК, а также σ.

Оценить адекватность построенной регрессионной модели с помощью критерия однородности.

2 Теория

2.1 Нахождение точечной оценки (β)

Выберем базис в $C^{(0)}[t_1, t_n]$ из полиномов и построим $\hat{g}(t)$ как их композицию:

$$\hat{g}(t) = \hat{\beta}_0 t^0 + \hat{\beta}_1 t^1 + ... + \hat{\beta}_k t^k \tag{1}$$

$$X(t) = z^T \beta + \xi(t) \tag{2}$$

, где

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \dots \\ \hat{\beta}_k \end{pmatrix}, t = \begin{pmatrix} t^0 \\ \dots \\ t^k \end{pmatrix}, z = \begin{pmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_n \\ \vdots & \vdots & \ddots & \vdots \\ t_1^k & t_2^k & \dots & t_n^k \end{pmatrix}, \xi(t) = \begin{pmatrix} \xi(t_1) \\ \dots \\ \xi(t_n) \end{pmatrix}$$

Чтобы найти точечные оценки $\hat{\beta}$ минимизируем функционал:

$$S(\hat{\beta}) := \sum_{i=1}^{n} (X(t_i) - \hat{g}(t_i))^2 \tag{3}$$

Из условия сущесвования глобального оптимума для выпуклой задачи: $\frac{\partial S}{\partial \beta} = 0$, сводим нахождение оценок к решению СЛАУ:

$$z * z^T * \hat{\beta} = z * X \tag{4}$$

2.1.1 Псевдообратная матрица

Теорема. Матрица $A^+ = V^T D^{-1} U$ является псевдообратной к матрице А. Док-во - просто рассмотреть AA^+. Тогда решение нормального уравнения сводится к

$$\hat{\beta} = (z * z^t)^+ * z * X$$

2.2 SVD разложение

Для того, чтобы решить СЛАУ, воспользуемся сингулярным разложением матрицы.

Определение 2.1 (SVD-разложение) *Представление вещесвенной $m \times n$ - матрицы $\boldsymbol{A}$ в виде:*

$$A = U\Sigma V^T \tag{5}$$

, где Σ - диагональная $m \times n$ - матрица с диагональю невозрастающих сингулярных чисел $\sigma_1, \ldots, \sigma_k$, а $\boldsymbol{U}$ и $\boldsymbol{V}$ - ортогональные, соответсвенно $m \times m$ и $n \times n$ - матрицы, называется ***сингулярным разложением*** *матрицы $\boldsymbol{A}$.*

Поскольку фигурирующие в сингулярном разложении матрицы $\boldsymbol{U}$ и $\boldsymbol{V}$ - ортогональные, вполне естесвенно получить их в результате выполнения последовательностей простых ортогональных преобразований.

Весь процесс сингулярного разложения можно условно разбить на два этапа. Первый этап состоит в приведении исходной $m \times n$ - матрицы $\boldsymbol{A}$ к двухдиагональной форме, что реализуется с помощью преобразований Хаусхолдера. На втором этапе двухдиагональная матрица - итог первого этапа - приводится к диагональному виду, что достигается применением к ней QR-алгоритма на основе преобразований Гивенса.

2.2.1 Этап двухдигонализации

Данный этап состоит в том, что матрица $\boldsymbol{A}$ умножается слева на матрицу Хаусхолдера $\boldsymbol{P_1}$, определяемую элементами a_{i1} первого столбца $\boldsymbol{A}$ так, чтобы аннулировать все его элементы, находящихся под диагональю.

Справа матрица $\boldsymbol{P_1 A}$ умножается тоже на матрицу Хаусхолдера $\boldsymbol{Q_1}$, определяемую элементами $\tilde{a}_{1j}$ первой строки матрицы $\boldsymbol{P_1 A}$ так, чтобы аннулировать все элементы, стоящие правее элемента $\tilde{a}_{12}$. Следовательно результат этого шага - матрица:

$$\boldsymbol{A_1 = P_1 A Q_1^T} \tag{6}$$

, где

$$P_1 = E - w_1 * w_1^T, w_1 = \mu_1 * (a_{11} - p_1; a_{21}; \ldots; a_{m1})^T$$

$$p_1 = sign(-a_{11})\sqrt{\sum_{i=1}^{m} a_{i1}^2}, \mu_1 = \frac{1}{\sqrt{2p_1^2 - 2 * p_1 * a_{11}}}$$

$$Q_1 = E - z_1 * z_1^T, z_1 = \nu_1 * (0, \tilde{a}_{12} - q_2; \tilde{a}_{13}; \ldots; \tilde{a}_{1n})^T$$

$$q_2 = sign(-\tilde{a}_{12})\sqrt{\sum_{j=1}^{m} \tilde{a}_{ij}^2}, \nu_2 = \frac{1}{\sqrt{2q_2^2 - 2 * q_2 * \tilde{a}_{12}}}$$

Нетредно представить, что в результате конечного процесса таких преобразований данная $m \times n$ - матрица $\boldsymbol{A}$ будет ортогональна приведена к двухдиагональной форме.

В нашем случае матрица квадратная ($m = n$), следовательно итог будет иметь вид:

$$\boldsymbol{B} = \begin{pmatrix} p_1 & q_2 & 0 & 0 & \ldots & 0 & 0 \\ 0 & p_2 & q_3 & 0 & \ldots & 0 & 0 \\ 0 & 0 & p_3 & q_4 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & p_{n-1} & q_n \\ 0 & 0 & 0 & 0 & \ldots & 0 & p_n \end{pmatrix} \tag{7}$$

$$\boldsymbol{A = PBQ^T}$$

2.2.2 Этап диагонализации матрицы B

Полученную двухдиагональную матрицу $\boldsymbol{B}$ приведем к диагональному виду с помощью ортогональных преобразований вида:

$$B_{n+1} = SB_nT^T \tag{8}$$

, которые совершаются на основе плоских вращений Гивенса. При этом матрицы $\boldsymbol{T}$, получаемые как произведения $n-1$ матриц вращений $\boldsymbol{T_2, T_3, \ldots, T_n}$, получаемых так, чтобы обеспечить сходимость последовательностей трехдиагональных матриц типа $\boldsymbol{F = B^T B}$ к диагональному виду в процессе QR-алгоритма.

На матрицы $\boldsymbol{S}$, так же представляющие собой произведения $n-1$ матриц Гивенса $\boldsymbol{S_2, S_3, \ldots, S_n}$, возлагается обязанность поддерживать двухдиагональную структуру матрицам типа матриц $\boldsymbol{B}$ на каждом полном промежудочном шаге алгоритма. Итоговое разложение матрицы $\boldsymbol{A}$:

$$\boldsymbol{A = PBQ^T = P(SDT^T)Q^T = UDV^T} \tag{9}$$

2.3 Нахождение точечной оценки σ

Любая точечная оценка, вычисленная на основании опытных данных, является их функцией и поэтому сама должна представлять собой случайную величину с распределением, зависящим от распределения исходной случайной величины, в том числе от самого оцениваемого параметра и от числа опытов $\boldsymbol{n}$.

К точечным оценкам предъявляется ряд требований, определяющих их пригодность для описания самих параметров.

1. Оценка называется состоятельной, если при увеличении числа наблюдений она приближается (сходится по вероятности) к значению оцениваемого параметра.

2. Оценка называется несмещенной, если ее математическое ожидание равно оцениваемому параметру.

3. Оценка называется эффективной, если ее дисперсия меньше дисперсии любой другой оценки данного параметра.

$$L(\xi_1, \xi_2, \ldots, \xi_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X(t_i)-g(t_i))^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi\sigma^2})^{\frac{n}{2}}} e^{-\frac{\sum_{i=1}^{n}(X(t_i)-g(t_i))^2}{2\sigma^2}}$$

$$\ln(L) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X(t_i) - g(t_i))^2 - n * \ln(\sigma) - \frac{n}{2}\ln(2\pi)$$

$$\left.\frac{d(\ln(L))}{d\sigma}\right|_{g=\hat{g}, \sigma=\hat{\sigma}} = \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n}(X_i - \hat{g}_i)^2 - \frac{n}{\hat{\sigma}} = 0$$

Откуда получаем:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{g})^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\xi}_t^2 \tag{10}$$

Но это смещенная оценка. Несмещенная выводится следующим образом:

$$M[S(\hat{\beta})] = M[(X - z^T\hat{\beta})^T(X - z^T\hat{\beta})] = \ldots = n\sigma^2 - k\sigma^2 = (n-k)\sigma^2$$

$$\hat{\sigma}^2 = \frac{1}{n-k}\sum_{i=1}^{n}(X_i - \hat{g})^2 = \frac{1}{n-k}\sum_{i=1}^{n}\hat{\xi}_t^2 \tag{11}$$

2.4 Проверка однородности выборки

Критерий Манна — Уитни — статистический критерий, используемый для оценки различий между двумя независимыми выборками по уровню какого-либо признака, измеренного количественно. Позволяет выявлять различия в значении параметра между малыми выборками.

Этот метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами (ранжированным рядом значений параметра в первой выборке и таким же во второй выборке). Чем меньше значение критерия, тем вероятнее, что различия между значениями параметра в выборках достоверны.

3 Реализация

Моделируем случайный процесс, задаваемый в виде: $X(t) = g(t) + \xi(t)$, где
$g(t)$ - детерменированая величина, задаваемая полиномом 3-ей степени (с разными доминирующими слагаемыми); $\xi(t)$ - случайная функция $\sim N(0, \sigma)$.

Задаем колличесво интервалов разбиения. На каждом интервале разбиения находим параметры регрессионной модели $\hat{\beta}$ и $\hat{\sigma}$. Находим ξ в каждой точке времени. Разделяем выборку на две части и оцениваем полученное P_{value}.

4 Результаты

Размер интервала разбиения для всех исследований брался равным 50 эллементов. Случайный процесс строился мощностью в 10^7 элементов.

1. При моделировании случайного процесса $g(t)$ было заданно с доминирующим слагаемым при 3-ей степени полинома:

   $$g(t) = 8t^3 + 0.04t^2 + 0.07t + 0.06$$

   Полученные результаты: ближе к началу временного интервала все модели дают хорошее приближение с точки зрения критерия однородности. Оценка $\hat{\sigma}$ приближенно равна заданной.

   Затем по критерию однородности не проходит модель с доминирующим совбодным членом. Потом перестает проходить модель с доминирующим слагаемым при первой степени (линейная). Потом с доминирующей второй и третьей.

2. $g(t)$ заданно с доминирующим слагаемым при 2-ей степени полинома:

   $$g(t) = 0.04t^3 + 8t^2 + 0.07t + 0.06$$

   Полученные результаты: результаты схожи с предыдущим пунктом.

3. $g(t)$ заданно с доминирующим слагаемым при 1-ей степени полинома:

   $$g(t) = 0.04t^3 + 0.07t^2 + 8t + 0.06$$

   Полученные результаты: результаты схожи с предыдущим пунктом.

4. $g(t)$ заданно с доминирующим слагаемым при 0-ей степени полинома:

   $$g(t) = 0.04t^3 + 0.07t^2 + 0.06t + 8$$

   Полученные результаты: результаты схожи с предыдущим пунктом. Модель с доминирующим совбодным членом работает лишь в окресности 0.

5 Выводы

Регрессионная модель, линейная по параметрам, адекватла лишь в окресности нуля. Чем дальше, тем хуже апроксимация тренда. Модели с доминирующими свободным челеном и коэффициентом при первой степени полинома раньше дают большую погрешность приближения, чем модели с доминирующими коэффициентами при 2 и 3 степенях полинома.

Точечная оценка σ растет во времени.