Параметрические методы идентификация линейных объектов

Гончаров Олег Игоревич

Факультет вычислительной математики и кибернетики
Московский государственный университет имени М.В. Ломоносова

Москва

Общий вид одношагового предсказателя

Строим предсказатель для объекта:

$$y[lT] = \sum_{k=1}^{\infty} g[kT]u[(l-k)T] + \sum_{k=0}^{\infty} h[kT]e[(l-k)T] \text{ или}$$

$$y(t) = G(E)u(t) + H(E)e(t),$$

где $H(E)$, $G(E)$ устойчивы, $h(0) = 1$, $u(t)$ — известный вход, $e(t)$ — белый шум с моментами $(0, \lambda)$.

Предсказатель:

$$\hat{v}(t|t-1) = (1 - H(E)^{-1})(y(t) - G(E)u(t)),$$
$$\hat{y}(t|t-1) = H(E)^{-1}G(E)u(t) + (1 - H(E)^{-1})y(t)$$

или

$$H(E)\hat{y}(t|t-1) = G(E)u(t) + (H(E) - 1)y(t).$$

Линейные регрессии

ARX-модель

$$A(E)y(t) = B(E)u(t) + e(t),$$

где $A(E) = 1 + a_1 E^{-1} + \dots + a_n E^{-n}$, $B(E) = b_0 + \dots + b_m E^{-m}$ — фильтры с конечной импульсной характеристикой, $e(t)$ — белый шум.

Обозначим $\theta = [a_1, a_2, \dots a_n, b_1, \dots b_n]^T$ — вектор параметров модели.

Предсказатель:

$$\hat{y}(t|\theta) = B(E)u(t) + [1 - A(E)]y(t) = \varphi^T(T)\theta,$$

где $\varphi(t) = [-y(t-1), \dots - y(t-n), u(t-1), \dots u(t-m)]^T$ — регрессионный вектор.

Псевдолинейная регрессия

ARMAX-модель

$$A(E)y(t) = B(E)u(t) + C(E)e(t),$$

где $C(E) = 1 + c_1 E^{-1} + \dots + c_k E^{-n}$.

Обозначим $\theta = [a_1, a_2, \dots a_n, b_0, \dots b_n, c_1, \dots c_k]^T$ — вектор параметров модели.

Предсказатель:

$$\hat{y}(t|\theta) = B(E)u(t) + [1 - A(E)]y(t) + [C(E) - 1][y(t) - \hat{y}(t|\theta)] = \varphi^T(t, \theta)\theta,$$

где $\varphi(t) = [-y(t-1), \dots u(t-1), \dots u(t-m), \varepsilon(t-1), \dots \varepsilon(t-k)]^T$ — регрессионный вектор ($\varepsilon(t) = y(t) - \hat{y}(t|\theta)$ — ошибка предсказания).

Параметрические методы оценки

Основная идея: “подгонка” предсказателя заданной структуры под имеющиеся данные $Z^N = (u^N, y^N)^T$.

- Сходимость: оценкам параметров предсказателя $\hat{\theta}^N$ сходятся при $N \to \infty$.
- Состоятельность: оценка $\hat{\theta}^N$ сходится к “истинным” параметрам объекта θ_0.
  - Истинный объект должен содержаться в классе рассматриваемых модельных структур.
  - Входные данные должны быть достаточно информативны.

Основные подходы параметрической идентификации:

- Миминизация ошибки предсказателя: МНК, метод максимального правдоподобия, информационные критерии оценки.
- Корреляционные методы: ошибка предсказания не должна коррелировать со входом.

Модельные структуры

Предсказатель задается парой фильтров $W_y(E)$ и $W_u(E)$:

$$\hat{y}(t) = W_y(E)y(t) + W_u(E)u(t).$$

Определение

Модельной структурой $\mathcal{M}$ называют отображение из множества параметров $\theta \in D_{\mathcal{M}} \subset \mathbb{R}^d$ во множество моделей $\mathcal{M} = \{W = [\, W_y(z) \; W_u(z) \,]\}$ такое, что градиенты функций предсказателя по θ устойчивы.

Замечание: для объектов вида $Ay = \frac{B}{F}u + \frac{C}{D}e$ достаточна устойчивость FIR фильтров F и C при любых θ.

Строгая устойчивость: равномерная устойчивость всего семейства фильтров предсказателя, их первых и вторых производных по θ.
Идентифицируемость: из совпадения АФЧХ моделей следует совпадение их параметров.

Информативность данных

Определение

Последовательность данных Z^∞ информативна относительно семейства моделей $\mathcal{M}^*$, если из $\bar{E}[(W_1(E) - W_2(E))z(t)]^2 = 0$ следует совпадение АФЧХ моделей $W_1(E)$ и $W_2(E)$.

Т.е. данные Z^∞ позволяют различать модели.

- Идентификация в открытом контуре:
  - достаточным условием является постоянство возбуждения порядка n ($n = \deg(B) + \deg(F)$) для сигнала $u(t)$, по определению такой сигнал не обнуляемый любым фильтром порядка не выше n.
- Идентификация в замкнутом контуре: задаем $w(t)$, а $u(t) = R(E)y(t) + w(t)$,
  - возникают существенные сложности: $y(t) + ay(t-1) = bu(t-1) + e(t)$, $u(t) = ky(t)$,
  - требуется накладывать ограничения на ОС: она не может быть стационарной,
  - проблему можно решить переключая несколько регуляторов.

Минимизация ошибки предсказателя

$$V_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} l\left(L(E)[y(t) - \hat{y}(t|\theta)]\right),$$
$$\hat{\theta}^N = \operatorname*{argmin}_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N).$$

где $l(\varepsilon)$ — функция нормы, $L(E)$ — предварительный фильтр, $\hat{y}(t|\theta)$ — предсказание.

- Функция $l(\varepsilon)$ дает разные методы: МНК, взвешенный МНК, нестационарный МНК, ММП и т.п.
- Предварительный фильтр $L(E)$ позволяет выделить значимые частоты, эффект эквивалентен домножению модели шума на $L(E)^{-1}$.
- Идентификация сводится к задаче оптимизации: градиентные методы, метод Ньютона.
- Существуют рекуррентные методы оценки, работающие online.

МНК и частотная интерпретация

В простейшем случае метода наименьших квадратов:

$$\text{Пусть } L(E) = 1, \; l(\varepsilon) = \frac{1}{2}\varepsilon^2, \text{ тогда}$$

$$\hat{\theta}^N = \operatorname*{argmin}_{\theta \in D_{\mathcal{M}}} \left[\frac{1}{N} \sum_{t=1}^{N} \frac{1}{2} (\hat{y}(t|\theta) - y(t))^2 \right].$$

возможна частотная интерпретация:

$$V_N(\theta, t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2} |\hat{\hat{G}}(e^{j\omega}) - G(e^{j\omega}, \theta)|^2 Q_N(\varepsilon^{j\omega}) d\omega,$$

где

$$Q_N(e^{j\omega}) = \frac{|U_N(\omega)|^2}{|H(e^{j\omega}, \theta)|^2}.$$

Т.е. происходит подгонка модели к эмпирической оценки передаточной функции с весом, определяемым модельным соотношением сигнал/шум.

Оценка параметров линейной регрессии

Применим МНК к предсказателю в форме линейной регрессии:

$$\hat{y}(t) = \varphi(t)^T \theta.$$

Задача сводится к нахождению псевдорешения переопределенной СЛАУ, тогда

$$\hat{\theta}_N^{LS} = \left[\frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varphi(t)^T \right]^{-1} \frac{1}{N} \sum_{t=1}^{N} \varphi(t) y(t) = R(N)^{-1} f(N).$$

В силу эргодических теорем $R(N) \to R^*$ и

$$\hat{\theta}_N \to \theta_0 + [R^*]^{-1} h^*, \text{ где } h^* = \bar{E}\varphi(t) v_0(t),$$

$v_0(t)$ — возмущение, действующее в истинной системе:
$y(t) = \varphi(t)^T \theta_0 + v_0(t)$.

Оценка параметров линейной регрессии

Условия состоятельности и сходимости:

- R^* — невырождена. Достаточным условием является постоянство возбуждения.
- $h^* = 0$. Возможно два случая:
  - $v(t)$ — белый шум,
  - $u(t)$ и $v(t)$ независимы (открытый контур) и фильтра $A(E) = 1$.

Т.е. ARX модель **не работает с цветным шумом**.

Это можно обойти в МНК с повторениями для ARX модели вида

$$A(E)D(E)y(t) = B(E)D(E)u(t) + e(t),$$

где $e(t)$ — белый шум, что эквивалентно модели возмущения

$$v(t) = \frac{1}{D(E)}e(t).$$

Корреляционный метод

Основная идея: ошибка предсказания не должны коррелировать с полученной информацией о системе.

Вычисляем ошибку предсказания

$$\varepsilon_f(t) = L(E)\varepsilon(t) = y(t) - \hat{y}(t|\theta),$$

вводим корреляционный вектор

$$\zeta(t,\theta) = \zeta(t, Z^{t-1}, \theta),$$

оценку параметров дает решение уравнения

$$\hat{\theta}_N = \mathrm{sol}\left[\frac{1}{N}\sum_{t=1}^{N}\zeta(t,\theta)\varepsilon_f(t,\theta) = 0\right].$$

Свобода состоит в выборе корреляционного вектора: он произвольный, но **не должен зависеть от значения выхода в предсказываемый момент** $y(t)$.

Псевдолинейная регрессия

Предсказатель имеет вид:

$$\hat{y}(t|\theta) = \varphi(t, \theta)^T \theta.$$

Используем вектор регрессоров в качестве корреляционного вектора:

$$\hat{\theta}_N^{PLR} = \underset{\theta \in D_{\mathcal{M}}}{\mathrm{sol}} \left[\frac{1}{N} \sum_{t=1}^{N} \varphi(t, \theta)[y(t) - \varphi(t, \theta)^T \theta] = 0 \right].$$

Метод инструментальных переменных

Предсказатель имеет вид линейной регрессии для ARX модели, но возмущение не является белым шумом

$$\hat{y}(t|\theta) = \varphi(t,\theta)^T \theta.$$

Введем корреляционный вектор (вектор инструментальных переменных):

$$\zeta(t) = K(E)\left[-x(t-1) \quad \dots \quad -x(t-n_a) \quad u(t-1) \quad \dots \quad u(t-n_b)\right],$$

$$\text{где } x(t) = \frac{M(E)}{N(E)} u(t), \quad M(E), N(E) \text{ — FIR фильтры.}$$

Общая практика состоит в выборе в качестве $M(E)$ и $N(E)$ оценок, полученных для $B(E)$ и $A(E)$ при помощи МНК.

$$\hat{\theta}_N^{IV} = \left[\frac{1}{N}\sum_{t=1}^{N} \zeta(t)\varphi(t)^T\right]^{-1} \frac{1}{N}\sum_{t=1}^{N} \zeta(t) y(t).$$

Замечание: при $\zeta(t) = \varphi(t)$ получаем МНК для линейной регрессии.

Сходимость

Пусть

- эксперимент проводится в открытом контуре или в замкнутом регулятором с дробно-рациональной ПФ,
- семейство моделей $\mathcal{M}$ равномерно устойчиво,

тогда имеет место сходимость последовательности оценок $\hat{\theta}_N$:

- оценка $\hat{\theta}_N$ сходится к множеству лучших приближений $D_c = \operatorname*{argmin}_{\theta \in D_{\mathcal{M}}} \bar{V}(\theta)$ для заданной последовательности данных Z^∞ и критерия $\bar{V}(\theta)$,
- в случае идентифицируемости семейства моделей $\mathcal{M}$ оценка $\hat{\theta}_N \to \theta^*$, где $\mathcal{M}(\theta^*)$ — лучшее приближение,
- с.в. $\sqrt{N}(\hat{\theta}_N - \theta_*)$ имеет гауссовское распределение,
- сходимость имеет место и в частотной области (см. частотную трактовку МНК).

Состоятельность

Пусть

- эксперимент проводится в открытом контуре или в замкнутом регулятором с дробно-рациональной ПФ,
- семейство моделей $\mathcal{M}$ равномерно устойчиво,
- истинная модель $\mathcal{M}_0$ принадлежит семейству $\mathcal{M}_0 = \mathcal{M}(\theta_0) \in \mathcal{M}$,
- последовательность данных Z^∞ вполне информативна,

тогда полученные оценки будут состоятельны:

- лучшее приближение совпадает с истинной моделью с точность до АФЧХ (равенства моделей),
- в случае идентифицируемости семейства моделей $\mathcal{M}$) имеем $\theta^* = \theta_0$, где $\mathcal{M}(\theta_0) = \mathcal{M}_0$ — истинной модели.