Нелинейная регрессия

$\exists$ $f(x, \alpha)$ - модель регрессии, Q - функционал качества, min по $\alpha \in \mathbb{R}^m$

$$Q(\alpha, X^l) = \sum_{i=1}^{l} (f(x_i,\alpha) - y_i)^2$$ ~~[illegible]~~

Для минимизации воспользуемся методом Ньютона - Рафсона.

$\alpha^0 = (\alpha_1^0, \alpha_2^0, \dots, \alpha_m^0)$ - начальное приближение

$\alpha^{t+1} = \alpha^t - h_t (Q''(\alpha^t))^{-1} Q'(\alpha^t)$ - итерационный процесс.

$Q'(\alpha^t)$ - градиент, $Q''(\alpha^t)$ - гессиан (матрица вторых производных) функционала Q в α^t, h - величина шага, который можно регулировать (либо в прост. случае $h = 1$)

$$\frac{\partial}{\partial \alpha_j} Q(\alpha) = 2 \sum_{i=1}^{l} \left(f(x_i,\alpha) - y_i\right) \frac{\partial f}{\partial \alpha_j}(x_i,\alpha)$$

$$\frac{\partial^2}{\partial \alpha_j \partial \alpha_k} Q(\alpha) = 2 \cdot \sum_{i=1}^{l} \frac{\partial f}{\partial \alpha_j}(x_i,\alpha) \frac{\partial f}{\partial \alpha_k}(x_i,\alpha) - 2 \sum_{i=1}^{l} (f(x_i,\alpha) - y_i) \frac{\partial^2 f}{\partial \alpha_j \partial \alpha_k}(x_i,\alpha)$$

Отметим.

$\exists$ ф-ция линейна в нек ε-окр т-ки α^t (для этого она должна быть непрерывно дифф), тогда

$$f(x_i,\alpha) = f(x_i,\alpha^t) + \sum_{j=1}^{m} \frac{\partial f}{\partial \alpha_j}(x_i,\alpha_j)(\alpha_j - \alpha_j^t)$$

Тогда, заменив в гессиане f на ее линеаризацию, 2-е слагаемое обнулится (Метод Ньютона - Гаусса)

Перепишем итерац ф-лу в матричном виде

$\exists F_t = \left(\frac{\partial f}{\partial \alpha_j}(x_i, \alpha^t)\right)_{i=1,l}^{j=1,m}$ - матрица первых произв $l \times m$ на t-й итерации.

$f_t = \left(f(x_i, \alpha^t)\right)_{i=1,l}$ - вектор знач аппрокс ф-ции на t-итер

Тогда для t-ой итерации метода Ньютона-Гаусса в матричном виде примет вид:

$$\alpha^{t+1} = \alpha^t - h_t \underbrace{(F_t^T F_t)^{-1} F_t^T (f^t - y_i)}_{\delta}$$

В правой 2 - решение ср задачи мн многомер регрессии.

$\|F_t \delta - (f^t - y)\|^2 \to \min$. Т.е. сводим обр нелин регр свелась к послед мин регр задач.

Нелин одномер преобр признаков

Если лин модель - необоснов для задачи и предложить адекватную нелин модель не удается, то можно постр лин модель над нелин.

$$f(x, \alpha) = \sum_{j=1}^{n} \varphi_j(f_j(x))$$

необх подобрать φ_j, при кот достигается мин. квадратичного функционала.

Надстройка с возвр

Основана на итерац уточнении φ_j. Вначале они полагаются линейными, $\varphi_j(x) = \alpha_j f_j(x)$, и неизв коэф α_j нах методами МНК. На каждом след шаге выб одна φ_j, все ост фиксируются и выб ф-ция строится заново

Для этого решается ст. задача наим кв.:

$$Q(\varphi_j, X^\ell) = \sum_{i=1}^{\ell} \Big(\varphi_j(f_j(x_i)) - \big(y_i - \sum_{k=1, k\ne j}^{n} \varphi_k(f_k(x_i))\big)\Big)^2 \to \min_{\varphi_j}$$

с обуч выборкой $Z^\ell = (f_j(x_i), z_i)_{i=1}^{\ell}$. Решается к одномер методом.

Обобщенные лин модели.

Пусть м-ль регр $f(x, \alpha)$ — линейная, но известна нелин ф-ция связи $g(f)$ между выходом модели f и целевым призн. y. З-ча аппроксимации:

$$Q(\alpha, X^\ell) = \sum_{i=1}^{\ell} \Big(g\big(\sum_{j=1}^{n} \alpha_j f_j(x_i)\big) - y_i\Big)^2 \to \min_{\alpha \in \mathbb{R}^n}$$

Линеаризуем $g(z)$

$$g(z) = g(z_i) + g'(z_i)(z - z_i)$$

Тогда Q аппроксимируется $\tilde{Q}$, квадратич по α:

$$\tilde{Q}(\alpha, X^\ell) = \sum_{i=1}^{\ell} \Big(g(z) + g'(z_i)\big(\sum_{j=1}^{n} \alpha_j f_j(x_i) - z_i\big) - y_i\Big)^2 =$$

$$= \sum_{i=1}^{\ell} \underbrace{(g'(z_i))^2}_{w_i} \Big(\sum_{j=1}^{n} \alpha_j f_j(x_i) - \underbrace{\big(z_i + \frac{y_i - g(z_i)}{g'(z_i)}\big)}_{\tilde{y}_i}\Big)^2 \to \min_{\alpha \in \mathbb{R}^n}$$

Некв. ф-ции потерь

Ф-ция Мешалкина:

$$\mathcal{L}(a, y) = 1 - \exp\Big(-\frac{1}{2\sigma}(a - y)^2\Big)$$

σ — дисперсия обычного шума, не связ с ~~обычными глм~~ большими выбросами.