Прорешиваем ебучую кр

28 декабря 2023 г.

1 Задача 2. Линейные регрессионные модели.

- Свести задачу к виду линейной регрессионной модели. Т.е. конвертировать условие задачи в формулу вида

$$X = Z\theta + \varepsilon;$$

  В которой:

  - X - это вектор результатов экспериментов/измерений/взвешиваний;
  - Z - это фиксированная матрица измерений (что и по сколько штук измеряли);
  - θ - это вектор неизвестных параметров - измеряемые величины/веса объектов;
  - ε - это вектор случайнго шума, погрешностей;

- Привести задачу к такому виду, что:

  - $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T : \ \forall i \ \varepsilon_i \sim N\left(0, \sigma^2\right)$ [Обычно из условия задачи сразу следует несмещенность];
  - Стобцы матрицы Z - ЛНЗ [Обычно сразу есть из условия задачи];
  - Элементы вектора ε - попарно некоррелированы [Обычно есть попарная независимость из условия задачи, откуда сразу же есть и некоррелированность];

  В таком случае наша модель будет называться **линейной гауссовской** моделью.

- Получив линейную гауссовскую модель можно воспользоваться двумя известными оценками, обладающими полезными свойствами:

  - $\hat{\theta} = \left(Z^T Z\right)^{-1} Z^T X$;
  - $\hat{\sigma^2} = \frac{1}{n-k}\|X - Z\hat{\theta}\|^2$. n - размерность X, k - размерность θ;

  Полученные оценки будут удовлетворять всем требованиям в задаче, так как они являются оптимальными в линейной гауссовской модели.

2 Задача 4. Проверка линейных гипотез.

В гауссовской линейной модели можно рассматривать различные гипотезы вида:

$$H_0 : T\theta = \tau \ V.S. \ H_1 : T\theta \neq \tau.$$

Для проверки таких гипотез есть хороший критерий - F-критерий.

- Собственно, начнем с того, чтобы свести условие к такому виду. Обозначим $k = rk\ T$.
- $B = T\left(Z^T Z\right)^{-1} T^T$, тогда верно следующее:

$$\frac{\left(T\hat{\theta} - \tau\right)^T B^{-1} \left(T\hat{\theta} - \tau\right)}{\|X - Z\hat{\theta}\|^2} \cdot \frac{n-k}{m} \sim F_{m,n-k}$$

- Финальный штрих - выразить критерий. F-Критерий УЗ α:

$$R = \{\frac{\left(T\hat{\theta} - \tau\right)^T B^{-1} \left(T\hat{\theta} - \tau\right)}{\|X - Z\hat{\theta}\|^2} \cdot \frac{n-k}{m} > f_{1-\alpha}\},$$

  где f_p - p -квантиль распределения Фишера со степенями свободы $m, n-k$.

- Как вы могли заметить, формула выше - очень страшная, даже для запоминания, а про подсчет руками, при размере матрицы B больше 1 уже даже и речи быть не может. Так как на кр предлагается задача с размером $B = 2$, то можно воспользоваться следующим упрощением:
- **Альтернативная формула F-критерия**. Верна такая формула для F-критерия:

$$R = \{\frac{min_{\theta : T\theta = \tau}\|X - Z\theta\|^2 - \|X - Z\hat{\theta}\|^2}{\|X - Z\hat{\theta}\|^2} \cdot \frac{n-k}{m} > f_{1-alpha}\},$$

  тогда, если $m + 1 = k$, то у $T\theta = \tau$ решения будут вида $c \cdot \beta$, где c-какой-то вектор. Тогда решение задачи $min_{\theta : T\theta = \tau}\|X - Z\theta\|^2$, сводится к решению квадратного уравнения.