Лекция 11.

Математическое ожидание и дисперсия случайного вектора

Def. $E\vec{X} = \begin{pmatrix} EX_1 \\ \vdots \\ EX_n \end{pmatrix}$ - математическое ожидание случайного вектора

Def. Дисперсией или матрицей ковариаций называется $D\vec{X} = E((\vec{X} - E\vec{X})(\vec{X} - E\vec{X})^T)$, элементами которой $d_{ij} = \operatorname{cov}(X_i, X_j), d_{ii} = D(X_i)$

Свойства:

1. $E(A\vec{X}) = AE\vec{X}$
2. $E(\vec{X} + \vec{B}) = E\vec{X} + \vec{B}$
3. $D(A\vec{X}) = AD\vec{X}A^T$
4. $D(\vec{X} + \vec{B}) = D\vec{X}$

Уравнение общей регрессии

Пусть результат X зависит от k факторов $Z_1, \dots, Z_k$. Рассматриваем теоретическую модель линейной регрессии:

$X = \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_k Z_k + \varepsilon$, где $\vec{Z} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_k \end{pmatrix}$ - вектор факторов, $\vec{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$ - вектор коэффициентов регрессии

Пусть проведено $n \geq k$ экспериментов, $\vec{Z}^{(i)} = \begin{pmatrix} Z_1^{(i)} \\ \vdots \\ Z_k^{(i)} \end{pmatrix}$ - значения факторов при i-ом эксперименте,

$\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ - соответствующие значения результатов

Согласно модели:

$$\begin{cases} X_1 = \beta_1 Z_1^{(1)} + \beta_2 Z_2^{(1)} + \dots + \beta_k Z_k^{(1)} + \varepsilon_1 \\ X_2 = \beta_1 Z_1^{(2)} + \beta_2 Z_2^{(2)} + \dots + \beta_k Z_k^{(2)} + \varepsilon_2 \\ \dots \\ X_n = \beta_1 Z_1^{(n)} + \beta_2 Z_2^{(n)} + \dots + \beta_k Z_k^{(n)} + \varepsilon_n \end{cases}$$

Или в матричной форме: $\vec{X} = Z^T\vec{\beta} + \vec{\varepsilon}$, где $Z = \begin{pmatrix} Z_1^{(1)} & Z_1^{(2)} & \dots & Z_1^{(n)} \\ Z_2^{(1)} & Z_2^{(2)} & \dots & Z_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_k^{(1)} & Z_k^{(2)} & \dots & Z_k^{(n)} \end{pmatrix}$ - матрица плана, $\vec{\varepsilon}$ - вектор теоретических ошибок

Наша цель такова: по данной матрице плана Z и вектора результатов $\vec{X}$ дать оценки неизвестных параметров регрессии β_i и параметров распределения ошибки ε

Nota. Заметим, что у данной модели мы не теряем свободный член b_0, так как при необходимости можно считать, что первый фактор тождественен единицы. То есть первая строка матрица плана будет состоять из единиц

Метод наименьших квадратов и нормальные уравнения

Будем считать, что выполнено условие Cond.1, что ранг матрицы $\operatorname{rang} Z = k$, то есть все строки матрицы плана независимы

Введем матрицу $A = ZZ^T$. Ее свойства:

1. A - квадратная и симметричная
2. A - положительно определенная
3. $\exists B = \sqrt{A}$, то есть $B^2 = A$

Пусть $\vec{B} = \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix}$ - вектор оценок $\vec{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$

Тогда эмпирическая модель регрессии $\hat{\vec{X}} = Z^T\vec{B}$, $\vec{\varepsilon}_i = X_i - \hat{X}_i$ - экспериментальная ошибка, или $\vec{\hat{\varepsilon}} = \begin{pmatrix} \hat{\varepsilon}_1 \\ \vdots \\ \hat{\varepsilon}_n \end{pmatrix} = \vec{X} - Z^T\vec{B}$ - вектор экспериментальных ошибок

По методу наименьших квадратов подбираем $\vec{B}$ таким образом, чтобы $L(\vec{B}) = \sum\limits_{i=1}^{n} \hat{\varepsilon}_i^2 \longrightarrow \min$

$\sum\limits_{i=1}^{n} \hat{\varepsilon}_i^2 = \|\vec{\hat{\varepsilon}}\|^2 = \|\vec{X} - Z^T\vec{B}\|^2$ - квадрат расстояния от точки $\vec{X}$ до $Z^T\vec{B}$ в $\mathbb{R}^n$

$Z^T\vec{B}$ - точка линейного подпространства, порожденного векторами $Z^T\vec{t}$, где $\vec{t} \in \mathbb{R}^k$

Nota. Согласно Cond.1 размерность линейной оболочки, порожденной вектором $Z^T\vec{t}$, $\dim\langle Z^T\vec{t}\rangle = k$

Функция $L(\vec{B})$ минимальна, когда квадрат расстояния от точки $\vec{X}$ до данного подпространства минимален, тогда вектор $Z^T\vec{B}$ – проекция вектора $\vec{X}$ на это подпространство

Таким образом, вектор $\vec{X} - Z^T\vec{B}$ должен быть ортогонален данному подпространству, то есть скалярное произведение вектора $\vec{X}$ и всех векторов подпространства равно 0

$(Z^T\vec{t}, \vec{X} - Z^T\vec{B}) = (Z^T\vec{t})^T(\vec{X} - Z^T\vec{B}) = \vec{t}^T(Z^T)^T(\vec{X} - Z^T\vec{B}) = \vec{t}^T Z(\vec{X} - Z^T\vec{B}) = \vec{t}^T(Z\vec{X} - ZZ^T\vec{B}) = 0\ \forall \vec{t} \in \mathbb{R}^k$
Так как всем векторами подпространства ортогонален только нулевой вектор, то получаем, что $Z\vec{X} - ZZ^T\vec{B} = 0$ или $A\vec{B} = Z\vec{X}$ - нормальное уравнение (или система нормальных уравнений)
Так как по свойству 2 матрица A невырожденная, то существует обратная, получаем решение системы: $\vec{B} = A^{-1}Z\vec{X}$

Свойства оценок метода наименьших квадратов

Добавим еще одно важное условие Cond.2: теоретические ошибки ε_i - независимы и имеют одинаковое нормальное распределение $N(0, \sigma^2)$. То есть $E\vec{\varepsilon} = \vec{0}$, $D\vec{\varepsilon} = \sigma^2 E_n$ (ковариации равны нулю в силу независимости)
Свойства:

1. $\vec{B} - \vec{\beta} = A^{-1}Z\vec{\varepsilon}$

   $\vec{B} - \vec{\beta} = A^{-1}Z\vec{X} - \vec{\beta} = A^{-1}Z(Z^T\vec{\beta} + \vec{\varepsilon}) - \vec{\beta} = A^{-1}Z\vec{\varepsilon}$

2. $\vec{B}$ - несмещенная оценка для вектора $\vec{\beta}$

   $E(\vec{B} - \vec{\beta}) = E(A^{-1}Z\vec{\varepsilon}) = A^{-1}ZE\vec{\varepsilon} = 0 \Longrightarrow E\vec{B} = \vec{\beta}$

3. Матрица ковариаций $D\vec{B} = \sigma^2 A^{-1}$

   $D\vec{B} = D(\vec{B} - \vec{\beta}) = D(A^{-1}Z\vec{\varepsilon}) = A^{-1}ZD\vec{\varepsilon}A^{-1T}Z^T = A^{-1}Z\sigma^2 E_n A^{-1T}Z^T = \sigma^2 A^{-1}(ZZ^T)A^{-1} = \sigma^2 A^{-1}$

   Следствие: дисперсии оценок b_i можно выразить через σ^2 и коэффициенты матрицы A^{-1}: $Db_i = \sigma^2 (A^{-1})_{ii}$

Оценка дисперсии случайного члена

Обозначим $\hat{\sigma}^2 = \frac{1}{n}\sum\limits_{i=1}^{n}\hat{\varepsilon}_i^2$. Ясно, что $\hat{\sigma}^2$ - точечная оценка неизвестной дисперсии σ^2, однако она является смещенной оценкой

Пусть выполнены Cond.1 и Cond.2, тогда $\frac{n\hat{\sigma}^2}{\sigma^2} \in H_{n-k}$ и не зависит от $\vec{B}$

Так как $\frac{n\hat{\sigma}^2}{\sigma^2} \in H_{n-k}$, то $E\hat{\sigma}^2 = \frac{\sigma^2}{n}E\frac{n\hat{\sigma}^2}{\sigma^2} = \frac{\sigma^2}{n}(n-k) = \frac{n-k}{n}\sigma^2 < \sigma^2$ - смещенная вниз оценка

Тогда несмещенной оценкой будет $S^2 = \frac{n}{n-k}\hat{\sigma}^2 = \frac{1}{n-k}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$