Конспект по метрике к КР-2

Лиза

1) Модель множественной регрессии

- Скалярная форма:

  $Y_i = \beta_0 + \beta_1 X_{1_i} + \dots + \beta_k X_{k_i} + \varepsilon_i,\ i = 1..n$

  $e_i = Y_i = \hat{Y}_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1_i} - \dots - \hat{\beta}_k X_{k_i}$

- Матричная форма:

  $Y = \beta_0 I_n + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon$

  $Y = X\beta + \varepsilon$

В скалярном случае надо решать $k+1$ нормальное уравнение. Выпишем в дифференциальной форме?

2) $RSS(\hat{\beta}) \to \min$

$$RSS = \sum_i e_i^2 = e^T e = (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

$$\frac{dRSS}{d\hat{\beta}} = d\hat{\beta}^T (-X)^T (Y - X\hat{\beta}) + (Y - X\hat{\beta})^T (-X) d\hat{\beta} = -2(Y - X\hat{\beta})^T X d\hat{\beta}$$

$$X^T (Y - X\hat{\beta}) = 0 \Rightarrow \boxed{\hat{\beta}_{k\times 1} = (X^T X)^{-1} X^T Y}$$

3) Оценка ковариационной матрицы$\hat{\beta}$(y – вектор)

$\mathbb{E}(Ay) = A\mathbb{E}(y)$

$\mathrm{Var}(y+z) = \mathrm{Var}(y) + \mathrm{Var}(z) + 2Cov(y,z)$

$\mathrm{Var}(Ay) = A\,\mathrm{Var}(y) A^T$

$Cov(Ay, Bz) = ACov(y,z)B^T$

$$\mathrm{Var}(\hat{\beta}) = Var((X^T X)^{-1} X^T y) = (X^T X)^{-1} X^T\,\mathrm{Var}(y) X (X^T X)^{-1} \overset{\mathrm{Var}(y) = \sigma_\varepsilon^2 I}{=} \sigma^2 (X^T X)^{-1}$$

Дисперсия одной оценки (оценочная, не настоящая): $\hat{\sigma_\varepsilon}(X^T X)^{-1}_{jj}$??

4) $\hat{\beta}_\varepsilon^2 = \dfrac{RSS}{n-k}$ – несмещенная оценка дисперсии ошибок

k – число регрессоров (включая константу)

5) $\hat{\mathrm{Var}}\hat{\beta} = \hat{\sigma_\varepsilon}^2 (X^T X)^{-1}$ – несмещенная оценка ковариационной матрица

6) Теорема Гаусса-Маркова

Если:

1. $Y = X\beta + u$
2. Оценивается регрессия $\hat{Y} = X\hat{\beta}$ с помощью МНК
3. X могут быть случайными, $\mathbb{P}$(у матрицыXесть ЛЗ столбцы) $= 0$
4. $\mathbb{E}(u \mid X) = 0,\ \mathrm{Var}(u \mid X) = \sigma^2 I$
5. $n \geq k$

То:

1. $\hat{\beta}$ существует и является единственной с вероятностью 1
2. $\hat{\beta}$ линейна по y : $\hat{\beta} = (X^T X)^{-1} X^T y$
3. $\mathbb{E}(\hat{\beta} \,|\, X) = \beta \,/\, \mathbb{E}(\hat{\beta}) = \beta$
4. $\hat{\beta}$ эффективна в классе линейных по y и несмещенных оценок

6) Показатели качества подгонки регрессии

- $R^2 = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$ – не работает для регрессии без свободного члена
- $0 \le R^2 \le 1$

Пусть $y = Y - \bar{Y}$

$\hat{y} = \hat{Y} - \bar{\hat{Y}}$

- $R^2 = \dfrac{\hat{\mathrm{Var}}(\hat{Y})}{\hat{\mathrm{Var}}(Y)}$
- $R^2 \to \max\limits_{\hat{\beta}} \sim RSS \to \min\limits_{\hat{\beta}}$
- $C\hat{orr}(Y, \hat{Y}) = \dfrac{\left(\sum\limits_i y_i \cdot \hat{y}_i\right)^2}{\sum\limits_i y_i^2 \sum\limits_i \hat{y}_i^2} = \dfrac{\left(\sum\limits_i (\hat{y}_i + e_i)\hat{y}_i\right)^2}{\sum\limits_i y_i^2 \sum\limits_i \hat{y}_i^2} =$

Бабкен