1 Contrast

Definition: Let θ₁, ..., θₜ be a set of variables, either parameters or statistics, and a₁, ..., aₜ be known constants. The quantity \( \sum_{i=1}^{t} a_i \theta_i \) is a linear combination. It is called a contrast if \( \sum_{i=1}^{t} a_i = 0 \). Furthermore, two contrasts, \( \sum_{i=1}^{t} a_i \theta_i \) and \( \sum_{i=1}^{t} b_i \theta_i \), are orthogonal if \( \sum_{i=1}^{t} a_i b_i = 0 \).

2 Polynomial Contrast in Regression, S-L 7.1

Suppose we observe from the scatter plot something more than linear trend, say quadratic trend, then, we may want to consider model \( Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i \). However, the design matrix \( X \) will then become

\[
\begin{pmatrix}
1 & x_1^2 \\
1 & x_2^2 \\
\vdots & \vdots \\
1 & x_n^2
\end{pmatrix}
\]

This type of matrix is usually ill-conditioned.

Solution: Instead, we consider \( Y_i = \gamma_0 \phi_0(x_i) + \gamma_1 \phi_1(x_i) + \ldots + \gamma_k \phi_k(x_i) + \epsilon_i \), where \( \phi_r(x_i) \) is an \( r \)th degree polynomial in \( x_i \) \( (r = 0, 1, \ldots, k) \). Moreover, we have \( \sum_{i=1}^{n} \phi_r(x_i) \phi_s(x_i) = 0 \), \( \forall r, s, r \neq s \). Now the design matrix becomes,

\[
\begin{pmatrix}
1 & \phi_1(x_1) & \ldots & \phi_k(x_1) \\
1 & \phi_1(x_2) & \ldots & \phi_k(x_2) \\
\vdots & \vdots & \ddots & \vdots \\
1 & \phi_1(x_n) & \ldots & \phi_k(x_n)
\end{pmatrix}
\]

How to construct? \( \phi_0(x) = 1 \) and \( \phi_1(x) = 2(x - a_1) \). Then we have \( \phi_{r+1}(x) = 2(x - a_{r+1})\phi_r(x) - b_r \phi_{r-1}(x) \). \( a_{r+1} \) and \( b_r \) help to make the orthogonal equation hold.

Pros: Make the Inversion much easier.
Cons: Hard to interpret the coefficient.

Question: Will \( \hat{Y} \) change in this case?

How to test Linear model against Quadratic Model using Polynomial Regression?

3 Polynomial Contrast in Experimental Design, S-L 7.1

Suppose we have an experimental design such that the treatment factors can be ordered and equally-spaced. e.g. 20, 40, 60. Then we first transform them to \( x_i = i - \frac{1}{2}(n + 1), i = 1, 2, \ldots, n \).

Then use an established system of orthogonal polynomials to generate orthogonal contrasts. \( \phi_0(x) = 1, \phi_1(x) = \lambda_1 x, \phi_2(x) = \lambda_2(x^2 - \frac{1}{12}(n^2 - 1)) \) etc. R code is contr.poly(n)
Take $n = 3$ for example,

4 Weighted Least Square

$Y = X\beta + \epsilon$, suppose now $E(\epsilon) = 0$ but $\text{Var}(\epsilon) = \sigma^2\Sigma$, where $\sigma^2$ is unknown but $\Sigma$ is known. Then, instead of OLS, we should use WLS, where $\hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$.

Reason, Gauss-Markov Theorem. Denote OLS by $\tilde{\beta} = (X'X)^{-1}X'Y$.

5 Linear Regression Under constraints 3.8.1

Lagrange Multiplier: Suppose we want to

$$\minimize \quad f(x)$$

$$\text{constrained to} \quad g(x) = c$$

where $x$ is a vector here, define $\Lambda(x, \lambda) = f(x) - \lambda(g(x) - c)$, and optimize this problem w.r.t. $(x, \lambda)$, that is, find $\nabla\Lambda(x, \lambda) = 0$.

Back to regression, suppose now we know $A\beta = c$, where $A$ is a full-rank matrix $q \times p$. Then we need to minimize $L(\beta, \lambda) = \|Y - X\beta\|_2^2 - \lambda(A\beta - c)$.

Take derivative for $\beta$ and $\lambda$, we get,

$$\begin{cases} A\beta = c \\ -2X'Y + 2X'X\beta + A'\lambda = 0 \end{cases}$$

The solution is $\hat{\beta}_H = \hat{\beta} + (X'X)^{-1}A'[A(X'X)^{-1}A']^{-1}(c - A\hat{\beta})$.

Useful identity, $\|Y - \hat{Y}_H\|^2 = \|\hat{Y} - \hat{Y}_H\|^2 + \|Y - \hat{Y}\|^2$. Derivation:

6 Linear Hypothesis Testing in R

anova(lm1, lm2)

lht(lm1, A, c)