1.5.2. Mathematical formulation

1.5.2.1. The initial assumption

1.5.2.2. The best linear unbiased prediction

1.5.2.3. The empirical best linear unbiased predictor

Kazuhei Kouno

Initial assumption

- Output

$$g : \mathbb{R}^{n_{features}} \to \mathbb{R}$$

$$X \mapsto y = g(X)$$

- Gaussian process for ML

$$G(X) = \underline{f(X)^T \beta} + \underline{Z(X)}$$

線形回帰モデル

ゼロ平均ガウス過程

$$C(X, X') = \sigma^2 R(|X - X'|)$$

The best linear unbiased prediction

$$\hat{G}(X) = G\left(X \middle| y_1 = g(X_1), \dots, y_{n_{samples}} = g\left(X_{n_{samples}}\right)\right)$$

- 線形性

$$\hat{G}(X) \equiv a(X)^T y$$

- 不偏性

$$E\left[G(X) - \hat{G}(X)\right] = 0$$

- 最小二乗誤差

$$\hat{G}(X)^* = \arg\min_{\hat{G}(X)} E\left[\left(G(X) - \hat{G}(X)\right)^2\right]$$

⇒

$$a(X)^* = \arg\min_{a(X)} E[(G(X) - a(X)^T y)^2]$$

$$\text{s.t.} \quad E[G(X) - a(X)^T y] = 0$$

The best linear unbiased prediction

- 期待値

$$\mu_{\hat{Y}}(X) = f(X)^T \hat{\beta} + r(X)^T \gamma$$

- 分散

$$\sigma_{\hat{Y}}^2(X) = \sigma_Y^2 \left(1 - r(X)^T R^{-1} r(X) + u(X)^T (F^T R^{-1} F)^{-1} u(X)\right)$$

$$R_{ij} = R(|X_i - X_j|, \theta),\ i, j = 1, \ldots, m$$

$$r_i = R(|X - X_i|, \theta),\ i = 1, \ldots, m$$

$$F_{ij} = f_i(X_j),\ i = 1, \ldots, p,\ j = 1, \ldots, m$$

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y$$

$$\gamma = R^{-1}(Y - F\hat{\beta})$$

$$u(X) = F^T R^{-1} r(X) - f(X)$$

事前に分からない

モデルを経験的に選択
(EBLUP)