Multiple Corr. Coef.

Multiple Correlation Coefficient

$r = \text{Correlation}(Y, \hat{Y})$

$= \left(\sum_i (Y_i - \bar{Y})(\hat{Y}_i - \bar{\hat{Y}})\right) / \sqrt{(\sum_i (Y_i-\bar{Y})^2)(\sum_i (\hat{Y}_i - \bar{\hat{Y}})^2)}$

If $1 \in C(X)$, $\hat{Y} = PY \Rightarrow 1^T\hat{Y} = 1^T PY$

$\Rightarrow \bar{\hat{Y}} = \bar{Y}$ ①

$= (P1)^T Y$

$= 1^T Y$

$\bar{Y} 1 = \frac{11^T Y}{n} = \underbrace{P_{C(1)}}_{11^T/n} Y \in C(X)$ ②

So $R^2 = r^2 = \frac{(\sum_i (Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}))^2}{(\sum_i (Y_i - \bar{Y})^2)(\sum_i (\hat{Y}_i - \bar{Y})^2)} \begin{matrix} (\Delta_1) \\ (\Delta_2) \end{matrix}$

$(\Delta_1) = \left(\sum_i (Y_i - \hat{Y}_i + \hat{Y}_i - \bar{Y})(\hat{Y}_i - \bar{Y})\right)^2$

$= \left(\sum_i (\hat{Y}_i - \bar{Y})^2\right)^2$ Since

- $\sum_i (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y}) = \underbrace{e^T}_{\in C(X)^\perp} (\underbrace{PY - \bar{y}1}_{\in C(X)}) = 0$

$\Rightarrow R^2 = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2}{\sum_i (Y_i - \bar{Y})^2} = \frac{SSReg}{TSS}$.

Ex: Test $\beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0$

$F = \frac{SSE_0 - SSE}{SSE} \frac{n-p}{p-1}$

$= \frac{n-p}{p-1} \frac{TSS - SSR_0 - (TSS - SSR)}{TSS - SSReg}$

$= \frac{1 - R_0^2 - (1 - R^2)}{1 - R^2} = \frac{R^2 - R_0^2}{1 - R^2} \frac{n-p}{p-1}$

$R_0^2 = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2}{TSS} = 0$ Since $\hat{Y}_i^0 = \bar{Y}\ \forall i$.

$\Rightarrow F = \frac{R^2}{1-R^2} \frac{n-p}{p-1}$

Test Statistic for H_0:

$F = \frac{R^2}{1-R^2} \frac{n-p}{p-1} \sim F_{p-1, n-p}$ Or

$R^2 = \frac{(p-1)F}{(n-p)+(p-1)F}$, note:

$X \sim F_{d_1, d_2} \Rightarrow \frac{d_1 X / d_2}{1 + d_1 X / d_2} \sim B(\frac{d_1}{2}, \frac{d_2}{2})$

$\Rightarrow R^2 \sim B(\frac{p-1}{2}, \frac{n-p}{2})$ &

$E R^2 = \frac{p-1}{n-1}$ (Exercise 4C. P113)

Geometry of LS coef.

$\theta = EY = X\beta$

$= \sum_{i=1}^{p} \beta_i X_i \in C(X)$

How to interpret β_i & $\hat{\beta}_i$ geometrically?

Focus on β_k & $\hat{\beta}_k$.

Let $V = \mathcal{L}(X_1, \dots, X_{p-1})$

$\hat{X}_k = P(X_k | V)$

$X_k^\perp = X_k - \hat{X}_k$

$X_k^\perp$ provides info. other than those provided by $X_1, \dots, X_{p-1} \in V$.

Observe that

$\langle \theta, X_p^\perp \rangle$

$= \langle \sum_{i=1}^{p} \beta_i X_i, X_p^\perp \rangle$

$= \beta_p \langle X_p^\perp, X_p^\perp \rangle \Rightarrow$

$\beta_p = \frac{\langle \theta, X_p^\perp \rangle}{\|X_p^\perp\|_2^2}$, $\hat{\beta}_p = \frac{\langle Y, X_p^\perp \rangle}{\|X_p^\perp\|_2^2}$

$$\mathrm{Var}\,\hat{\beta}_p = \frac{1}{\|x_p^{\perp}\|_2^2}\mathrm{Var}(x_p^{\perp\prime}Y) = \frac{\sigma^2}{\|x_p^{\perp}\|_2^2}$$

$\therefore$ $\mathrm{Var}\,\hat{\beta}_p$ Large if $\|x_p^{\perp}\|$ is small

i.e. $x_p - \sum_{i=1}^{p-1}\beta_i x_i \approx 0$

i.e. if x_p is nearly a linear Combination of other x's. then $\hat{\beta}_p$ is very poorly estimated

$$\mathrm{Cov}(\hat{\beta}_i, \hat{\beta}_j) = \mathrm{Cov}\left(\frac{\langle x_i^{\perp}, Y\rangle}{\|x_i^{\perp}\|_2^2}, \frac{\langle x_j^{\perp}, Y\rangle}{\|x_j^{\perp}\|_2^2}\right)$$

$$= \frac{\sigma^2\langle x_i^{\perp}, x_j^{\perp}\rangle}{\|x_i^{\perp}\|_2^2\|x_j^{\perp}\|_2^2}$$

$$= \cos(x_i^{\perp}, x_j^{\perp})\frac{\sigma^2}{\|x_i\|_2\|x_j\|_2}$$

Thus, $\sigma^2(X^TX)^{-1}_{ij} = \frac{\sigma^2}{\|x_i\|_2\|x_j\|_2}\cos(x_i^{\perp}, x_j^{\perp})$

Relation to R^2:

$R^2 = SSReg/TSS$

$R^2_{k-1} = SSReg(x_1, \dots, x_{k-1})/TSS$

$R^2_k = SSReg(x_1, \dots, x_k)/TSS$

Verify $\frac{R_k^2 - R_{k-1}^2}{1-R_{k-1}^2} = \frac{t^2}{n-k+t^2}$,

t = test stat for $\beta_k = 0$.

Partial Correlation

Let $v_1, v_2, x_1, x_2, \dots, x_k \in \mathbb{R}^n$,

$V = \mathcal{L}(x_1, \dots, x_k)$

$\hat{v}_i = P(v_i|V), i=1,2$

$v_i^{\perp} = v_i - \hat{v}_i \in V^{\perp}, i=1,2$

The partial corr. coef. of V_1 & V_2 (with the linear effects of $x_1, \dots, x_k$ removed) is:

$$r_{v_1v_2\cdot x_1\cdots x_k} = \frac{\langle v_1^{\perp}, v_2^{\perp}\rangle}{\|v_1^{\perp}\|\|v_2^{\perp}\|}$$

$$\triangleq \cos_{x_1\cdots x_k}(v_1^{\perp}, v_2^{\perp})$$

NOTE:

- $k=1$: $x_1 = \mathbf{1}$, $v_1^{\perp} = v_1 - P(v_1|\mathbf{1}) = v_1 - \bar{v}_1\mathbf{1}$

$\Rightarrow$ $r_{v_1v_2\cdot 1}$ = raw corr. (pearson corr.)

Questions: Want to express pcc (partial corr. coef.) of order n in terms of pcc's of order $(n-1)$.

Ex $k=4$, x_1, x_2, x_3, x_4.

$$r_{14\cdot 23} = \frac{r_{14} - r_{13\cdot 2}\,r_{34\cdot 2}}{\sqrt{1-r_{13\cdot 2}^2}\sqrt{1-r_{34\cdot 2}^2}}$$

- Consider a general case:

$x_1, x_2, \dots, x_{k-1}, x_k$

& Let $V_J = \mathcal{L}(x_4, x_5, \dots, x_k)$

$\subseteq \mathcal{L}(x_3, x_4, \dots, x_k)$

$= V_I$, $I = \{3, \dots, k\}$, $J = \{4, \dots, k\}$

Let $\hat{x}_i = P(x_i|V_J)$ $i=1,2$

$x_i^{\perp} = x_i - \hat{x}_i \in V_J^{\perp}$, $i=1,2$

$x_3 = \hat{x}_3 + x_3^{\perp} = P(x_3|V_J) + (x_3 - P(x_3|V_J))$

For $i=1,2,\dots$, WT find

$p(x_i \mid V_I)$. To do this,

recall $(Q_1 = I - P_{C(X_1)})$

$$P_{C(X_1|X_2)} = P_{C(X_1)} + Q_1 X_2 (X_2^T Q_1 X_2)^{-1} X_2^T Q_1$$

(Do by orthogonization process:

$$EY = X_1\beta_1 + X_2\beta_2$$
$$= X_1(\beta_1 + (X_1^T X_1)^{-1} X_1^T X_2 \beta_2) + Q_1 X_2 \beta_2)$$

$$w_i = x_i - P(x_i \mid V_I) \quad \text{for } i=1,2.$$
$$= x_i - (p(x_i \mid V_J) + p(x_i \mid x_3^\perp))$$
$$= x_i^\perp - p(\underbrace{x_i - \hat{x}_i}_{①} + \underbrace{\hat{x}_i}_{②} \mid \underbrace{x_3^\perp}_{③})$$
$$= x_i^\perp - p(x_i^\perp \mid x_3^\perp)$$ since ② $\in V_J$, ③ $\in V_J^\perp$

So $r_{12\cdot I}$ = pcc of x_1 & x_2 adjusting for $x_3, x_4, \dots, x_k$

$$= \frac{\langle w_1, w_2 \rangle}{\|w_1\| \|w_2\|},$$ depends only on $x_1^\perp, x_2^\perp$ & $x_3^\perp$.

$$\|w_1\|^2 = \langle w_1, w_1 \rangle = \langle x_1^\perp - p(x_1^\perp \mid x_3^\perp), x_1^\perp - p(x_1^\perp \mid x_3^\perp) \rangle$$
$$= \left\langle x_1^\perp - \frac{\langle x_1^\perp, x_3^\perp \rangle}{\|x_3^\perp\|_2^2} x_3^\perp, x_1^\perp - \frac{\langle x_1^\perp, x_3^\perp \rangle}{\|x_3^\perp\|_2^2} x_3^\perp \right\rangle$$
$$= \langle x_1^\perp, x_1^\perp \rangle - \frac{\langle x_1^\perp, x_3^\perp \rangle^2}{\|x_3^\perp\|_2^2}$$

WLOG assume $\|x_i^\perp\|_2 = 1,\ i=1,2,3$

$$= 1 - \langle x_1^\perp, x_3^\perp \rangle^2 / \|x_1^\perp\|_2 \|x_3^\perp\|_2$$
$$= 1 - r_{13\cdot J}^2.$$

Similarly,

- $\|w_2\|^2 = 1 - r_{23\cdot J}^2$
- $\langle w_1, w_2 \rangle = \langle x_1^\perp - p(x_1^\perp \mid x_3^\perp), x_2^\perp - p(x_2^\perp \mid x_3^\perp) \rangle$

$$= r_{12\cdot J} - r_{23\cdot J}\, r_{13\cdot J}$$

$(\|x_i^\perp\|_2 = 1,\ i=1,2,3)$

In summary we showed

$$r_{12\cdot I} = r_{12\cdot 34\cdots k}$$
$$= \frac{r_{12\cdot J} - r_{13\cdot J}\, r_{23\cdot J}}{\sqrt{1 - r_{13\cdot J}^2}\sqrt{1 - r_{23\cdot J}^2}}$$

LHS: I & RHS: J. $\square$.