STOCKHOLMS UNIVERSITET
MATEMATISK STATISTIK

Lösningar till tentamensskrivning för kursen Linjära statistiska modeller

25 oktober 2019 9–14

Examinator: Ola Hössjer, tel. 070/672 12 18, ola@math.su.se

Uppgift 1

a) Låt $j = 1, 2, 3, 4, 5$ beteckna gruppnummer, n_j antal personer och $\bar{Y}_j$ medelvärdet av Y_i för alla personer i grupp j. Eftersom interceptet α är centrerat följer att

$$\begin{aligned}
\hat{\alpha} &= \bar{Y} = \textstyle\sum_{i=1}^{30} Y_i/30 \\
&= (n_1\bar{Y}_1 + n_2\bar{Y}_2 + n_3\bar{Y}_3 + n_4\bar{Y}_4 + n_5\bar{Y}_5)/30 \\
&= (4 \cdot 19.0 + 6 \cdot 22.0 + 10 \cdot 24.0 + 6 \cdot 25.5 + 4 \cdot 26.5)/30 \\
&= 23.57.
\end{aligned}$$

Med hjälp av ledningen får vi direkt att skattningen av lutningsparametern β ges av

$$\begin{aligned}
\hat{\beta} &= \textstyle\sum_{i=1}^{30}(x_i - \bar{x})Y_i / \sum_{i=1}^{30}(x_i - \bar{x})^2 \\
&= 81/40 = 1.841,
\end{aligned}$$

där

$$\begin{aligned}
\bar{x} &= \textstyle\sum_{i=1}^{30} /30 \\
&= (n_1 \cdot 5 + n_2 \cdot 6 + n_3 \cdot 7 + n_4 \cdot 8 + n_5 \cdot 9)/30 \\
&= (4 \cdot 5 + 6 \cdot 6 + 10 \cdot 7 + 6 \cdot 8 + 4 \cdot 9)/30 \\
&= 7.
\end{aligned}$$

b) Låt $N = 30$ vara antalet observationer. Skattningsvektorn $(\hat{\alpha}, \hat{\beta})^T$ är tvådimensionellt normalfördelad med väntevärde $(\alpha, \beta)^T$ och kovariansmatris

$$\begin{pmatrix} \sigma^2/N & 0 \\ 0 & \sigma^2/\sum_{i=1}^{N}(x_i - \bar{x})^2 \end{pmatrix} = \begin{pmatrix} \sigma^2/30 & 0 \\ 0 & \sigma^2/44 \end{pmatrix}.$$

c) Eftersom den enkla linjära regressionsmodellen innehåller 2 parametrar är antalet frihetsgrader för variationskällan Residual lika med $30 - 2 = 28$. Det ger en väntevärdesriktig skattning

$$\hat{\sigma}^2 = \text{Mkvs(Residual)} = \frac{\text{Kvs(Residual)}}{28} = \frac{550}{28} = 19.64$$

av variansparametern σ^2.

d) En person som sovit 6.5 timmar har en förväntad minnesförmåga

$$\mu = \alpha + (6.5 - \bar{x})\beta = \alpha - 0.5\beta.$$

Motsvarande skattning

$$\hat{\mu} = \hat{\alpha} - 0.5\hat{\beta} = 23.57 - 0.5 \cdot 1.841 = 22.65$$

har variansen

$$\mathrm{Var}(\hat{\mu}) = \mathrm{Var}(\hat{\alpha}) + 0.25\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{30} + \frac{0.25\sigma^2}{44} = 0.039\sigma^2.$$

Det ger ett medelfel

$$d = \sqrt{\widehat{\mathrm{Var}}(\hat{\mu})} = \sqrt{0.039\hat{\sigma}^2} = \sqrt{0.039 \cdot 19.64} = 0.8753$$

för $\hat{\mu}$ och ett 95% konfidensintervall

$$\begin{aligned}(\hat{\mu} - t_{0.025}(28)d, \hat{\mu} + t_{0.025}(28)d) &= (22.65 - 2.0484 \cdot 0.8753, 22.65 + 2.0484 \cdot 0.8753) \\ &= (20.85, 24.44)\end{aligned}$$

för μ. (Här fås t-kvantilen ur tabell med F-fördelningens kvantiler genom $t_{0.025}(28) = \sqrt{F_{0.05}(1, 28)}$.)

Uppgift 2

a) Modellen kan skrivas som

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \tag{1}$$

för syreupptagningsförmågan hos person $k \in \{1, \ldots, 4\}$ inom gruppen för vilka rökning är på nivån $i \in \{1, 2, 3\}$ och den fysiska aktiviteten på nivån $j \in \{1, 2\}$. Vidare är μ det genomsnlittliga väntevärdet för alla grupper, α_i den systematiska effekten av rökning på nivå i, β_j den den systematiska effekten av fysisk aktivitet på nivån j samt γ_{ij} samspelet mellan rökning och fysisk aktivitet. För att undvika överparametrisering inför vi totalt 6 linjärt oberoende bivillkor $\sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$ (varav 1 bivillkor för α_i, 1 för β_j och 3+2-1=4 för γ_{ij}). Feltermerna $\varepsilon_i \sim N(0, \sigma^2)$ antas vara oberoende.

b) För att testa grundmodellen (1) mot hypotesmodellen

$$\gamma_{ij} = 0, \ \forall i, j$$

att det inte finns något samspel mellan rökning och fysisk aktivitet, bildar vi

$$\text{F-kvot} = \frac{\text{Mkvs(Samspel)}}{\text{Mkvs(Inom celler)}} = \frac{\text{Kvs(Samspel)}/2}{\text{Kvs(Inom celler)}/18} = \frac{5.5/2}{19.5/18} = 2.54.$$

Här utnyttjade vi att variationskällan Sampspel har $(2-1)(3-1) = 2$ frihetsgrader, medan Inom celler har $3 \cdot 2(4-1) = 18$ frihetsgrader. Då F-kvoten har en $F(2, 18)$-fördelning under H_0 så jämför vi dess observerade värde med

$$F_{0.05}(2, 18) = 3.55.$$

Eftersom F-kvoten inte överstiger detta värde kan vi inte förkasta H_0 på signifikansnivån 5%.

c) Eftersom samsplet i b) inte var signifikant så antar vi en additiv modell (=hypotesmodellen i b). Alltså slår vi ihop de två variationskällorna Samspel och Inom celler till en ny variationskälla med 2+18=20 frihetsgrader. Vi skattar sedan feltermernas varians enligt

$$\hat{\sigma}^2 = \frac{\text{Kvs(Samspel)} + \text{Kvs(Inom celler)}}{2+18} = \frac{5.5 + 19.5}{20} = 1.25.$$

Eftersom variationskällan Rökning har 3-1=2 frihetsgrader får vi en

$$\text{F-kvot} = \frac{\text{Kvs(Rökning)}/2}{\hat{\sigma}^2} = \frac{10.0/2}{1.25} = 4.0 > F_{0.05}(2, 20) = 3.48.$$

Således kan vi förkasta nollhypotesen att rökning inte har någon effekt på syreupptagningsförmågan, på nivån 5%.

Uppgift 3

a) Vi kompletterar teckenschemat för det första fraktionella försöket med kolumner för enheten I och alla interaktioner av ordning 2 och 3:

I	C	P	T	CP	CT	PT	CPT
+	+	-	-	-	-	+	+
+	-	+	-	-	+	-	+
+	-	-	+	+	-	-	+
+	+	+	+	+	+	+	+

Genom att para ihop kolumnerna får vi kopplingsmönstret $I = CPT$, $C = PT$, $P = CT$, $T = CP$.

För det andra fraktionella försöket gör vi på motsvarande sätt. Utfyllnad av teckentabellen ger

I	C	P	T	CP	CT	PT	CPT
+	-	-	-	+	+	+	-
+	+	-	+	-	+	-	-
+	-	+	-	-	+	-	+
+	+	+	+	+	+	+	+

Genom att identifiera kolumnerna parvis ser vi att kopplingsmönstret är $I = CT$, $C = T$, $P = CPT$, $CP = PT$. Alternativt kan vi först notera att

CT är kopplad till enheten I, och sedan bestämma de andra tre kopplingarna utifrån det, t ex $C = CI = C(CT) = C^2T = T$ osv.

b) I det andra fraktionella försöket är en av kopplingarna $P = T$, så dessa två huvudeffekter kan inte särskiljas. För det andra fraktionella försöket tillhör de tre huvudeffekterna olika par av kopplade effekter. Varje huvudeffekt är alltså kopplad till en interaktionseffekt. Eftersom alla interaktionseffekter satts till 0 kan alla tre huvudeffekterna $\bar{C}$, $\bar{P}$ och $\bar{T}$ skattas för detta försök.

För att skatta huvudeffekterna för det första fraktionella försöket inför vi observationsvektorn $\boldsymbol{Y} = (Y_{+--}, Y_{-+-}, Y_{--+}, Y_{+++})^T$, parametervektorn $\boldsymbol{\theta} = (\mu, \bar{C}, \bar{P}, \bar{T})^T$, och designmatrisen

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

som fås genom att till det givna teckenschemat addera en kolumn med ettor (svarande mot μ). Man kan sedan använda den allmänna formeln

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{Y} = \frac{1}{4}\boldsymbol{A}^T\boldsymbol{Y}$$

för minsta kvadrat-skattningen av $\boldsymbol{\theta}$. Efter lite räkningar ser man att skattningarna av de tre huvudeffekterna blir

$$\begin{aligned} \hat{C} &= (Y_{+--} - Y_{-+-} - Y_{--+} + Y_{+++})/4 = 0.75, \\ \hat{P} &= (-Y_{+--} + Y_{-+-} - Y_{--+} + Y_{+++})/4 = 1.25, \\ \hat{T} &= (-Y_{+--} - Y_{-+-} + Y_{--+} + Y_{+++})/4 = 2.25. \end{aligned} \tag{2}$$

Alternativt kan man komma fram till (2) direkt genom att utgå från det andra försökets teckenschema, eftersom dess kolumner är ortogonala.

c) Vi börjar med att bestämma kovariansmatrisen för skattningen av parametervektorn $\boldsymbol{\theta}$. Den ges av

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) = \sigma^2(\boldsymbol{A}^T\boldsymbol{A})^{-1} = \frac{\sigma^2}{4}\boldsymbol{I}_4,$$

där $\boldsymbol{I}_4$ är identitetsmatrisen av ordning 4. Vidare har vi att

$$\begin{aligned} \Delta &= \mu_{+++} - \mu_{---} \\ &= (\mu + \bar{C} + \bar{P} + \bar{T}) - (\mu - \bar{C} - \bar{P} - \bar{T}) \\ &= 2(\bar{C} + \bar{P} + \bar{T}) \\ &= \boldsymbol{c}^T\boldsymbol{\theta}, \end{aligned}$$

där $\boldsymbol{c} = (0, 2, 2, 2)^T$. Av detta följer att

$$\hat{\Delta} = \boldsymbol{c}^T\hat{\boldsymbol{\theta}} = 2(\hat{C} + \hat{P} + \hat{T}) = 2(0.75 + 1.25 + 2.25) = 8.5$$

och

$$\begin{aligned}
\mathrm{Var}(\hat{\Delta}) &= \boldsymbol{c}^T \mathrm{Var}(\hat{\boldsymbol{\theta}})\boldsymbol{c} \\
&= \sigma^2/4 \cdot \boldsymbol{c}^T\boldsymbol{c} \\
&= \sigma^2/4 \cdot \textstyle\sum_{i=1}^4 c_i^2 \\
&= \sigma^2/4 \cdot (0^2 + 2^2 + 2^2 + 2^2) \\
&= 3\sigma^2.
\end{aligned} \tag{3}$$

Eftersom antalet regressionsparametrar $k = 4$ är lika med antalet observationer N, blir alla residualer 0. Det finns därför inga frihetsgrader kvar att skatta σ^2. Därmed kan inte heller variansen i (3) skattas.

Uppgift 4

a) Den givna modellen (ekvation (3) i skrivningsbladet) kan skrivas på matrisform som $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, där

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} Z_1 - 0.5 \\ Z_2 - 0.5 \\ Z_3 - 0.5 \\ Z_4 - 0.5 \end{pmatrix} = \begin{pmatrix} -0.27 \\ -0.09 \\ 0.12 \\ 0.24 \end{pmatrix},$$

är observationsvektorn,

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ x_{13} & x_{23} \\ x_{14} & x_{24} \end{pmatrix} = \begin{pmatrix} -0.5 & -0.5 \\ 0.5 & -0.5 \\ -0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

är designmatrisen och $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)^T$ feltermsvektorn. Vi börjar med att räkna ut

$$\boldsymbol{S} = \boldsymbol{X}^T\boldsymbol{X} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Det ger en minsta-kvadratskattning

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \boldsymbol{S}^{-1}\boldsymbol{X}^T\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{Y} = \begin{pmatrix} 0.5(-Y_1 + Y_2 - Y_3 + Y_4) \\ 0.5(-Y_1 - Y_2 + Y_3 + Y_4) \end{pmatrix} = \begin{pmatrix} 0.15 \\ 0.36 \end{pmatrix}.$$

b) Kovariansmatrisen för $\hat{\boldsymbol{\beta}}$ ges av

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{S}^{-1} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

Variansinflationsfaktorn för $\hat{\beta}_1$ anger hur mycket variansen av skattningen av β_1 ökar på grund av att man även måste skatta β_2. Eftersom variansen för skattningen av β_1 är σ^2/s_{11} då β_2 är känd, och $\sigma^2(\boldsymbol{S}^{-1})_{11}$ då β_2 är okänd, följer att

$$\mathrm{VIF}(\hat{\beta}_1) = \frac{\sigma^2(\boldsymbol{S}^{-1})_{11}}{\sigma^2/s_{11}} = s_{11} \cdot (\boldsymbol{S}^{-1})_{11} = 1 \cdot 1 = 1.$$

Variansinflationsfaktorn är alltså 1 eftersom de två förklarande variablerna x_1 och x_2 är ortogonala.

c) Vi börjar med att skatta feltermsvariansen. Eftersom residualerna har $N - 2 = 4 - 2 = 2$ frihetsgrader följer av ledningen och de uträknade skattningarna av β_1 och β_2 i a), att

$$\begin{aligned}
\hat{\sigma}^2 &= \tfrac{1}{2}\textstyle\sum_{i=1}^4 (Y_i - \hat{\beta}_1 x_{1i} - \hat{\beta}_2 x_{2i})^2 \\
&= \tfrac{1}{2}\left(\textstyle\sum_{i=1}^4 Y_i^2 - \hat{\beta}_1^2 \sum_{i=1}^4 x_{1i}^2 - \hat{\beta}_2^2 \sum_{i=1}^4 x_{2i}^2\right) \\
&= \tfrac{1}{2}\left(\textstyle\sum_{i=1}^4 Y_i^2 - \hat{\beta}_1^2 s_{11} - \hat{\beta}_2^2 s_{22}\right) \\
&= \tfrac{1}{2}(0.153 - 0.15^2 - 0.36^2) \\
&= 4.5 \cdot 10^{-4}.
\end{aligned}$$

Låt $\boldsymbol{\mu} = E(\boldsymbol{Y}) = \boldsymbol{X}\boldsymbol{\beta}$ vara väntevärdesvektorn för observationerna. Varje värde på $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$ kan testas som en nollhypotes, baserat på en

$$\text{F-kvot} = \frac{\|\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2/2}{\hat{\sigma}^2} = \frac{\|\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|^2}{2\hat{\sigma}^2} = \frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T \boldsymbol{S}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{2\hat{\sigma}^2}.$$

Eftersom $\boldsymbol{S} = \boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{I}_2$ är enhetsmatrisen av ordning 2, enligt a), så följer att

$$\text{F-kvot} = \frac{(\hat{\beta}_1 - \beta_1)^2 + (\hat{\beta}_2 - \beta_2)^2}{2 \cdot 4.5 \cdot 10^{-4}}.$$

Denna F-kvot har en $F(2, 2)$-fördelning under nollhypotesen. Det ger en konfidensregion E med konfidensgrad 95% som består av alla värden på $(\beta_1, \beta_2)^T$ för vilka nollhypotesen inte förkastas, det vill säga de värden på $(\beta_1, \beta_2)^T$ för vilka F-kvoten ovan inte överstiger $F_{0.05}(2, 2) = 19.0$. Det ger

$$\begin{aligned}
E &= \{(\beta_1, \beta_2)^T;\ (\hat{\beta}_1 - \beta_1)^2 + (\hat{\beta}_2 - \beta_2)^2 \le 2 \cdot 4.5 \cdot 10^{-4} \cdot 19.0\} \\
&= \{(\beta_1, \beta_2)^T;\ (0.15 - \beta_1)^2 + (0.36 - \beta_2)^2 \le 0.172\}.
\end{aligned}$$

Uppgift 5

a) Förklaringsgraderna för grund- respektive hypotesmodellerna anger hur stor andel av variationen i responsvariablerna Y_i som fångas upp av $\hat{\mu}_i$ respektive $\hat{\hat{\mu}}_i$. Det svarar mot Kvs(Regression)/Kvs(Total) för respektive modell, dvs

$$R_0^2 = \frac{\sum_{i=1}^N (\hat{\mu}_i - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} = \frac{\|\hat{\boldsymbol{\mu}} - \bar{Y}\|^2}{\|\boldsymbol{Y} - \bar{Y}\|^2} \tag{4}$$

och

$$R_1^2 = \frac{\sum_{i=1}^N (\hat{\hat{\mu}}_i - \bar{Y})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} = \frac{\|\hat{\hat{\boldsymbol{\mu}}} - \bar{Y}\|^2}{\|\boldsymbol{Y} - \bar{Y}\|^2}. \tag{5}$$

Vi införde här observationsvektorn $\boldsymbol{Y} = (Y_1, \ldots, Y_N)^T$ och vektorn $\bar{Y} = (\bar{Y}, \ldots, \bar{Y})^T$ som har identiska koordinater lika med skattningen av interceptet (dvs $\hat{\alpha} = \bar{Y}$).

b) Vektorerna $\hat{\boldsymbol{\mu}}$ och $\hat{\hat{\boldsymbol{\mu}}}$ är projektioner av observationsvektorn $\boldsymbol{Y}$ ned på de delrum $\mathcal{U}_k$ och $\mathcal{U}_l \subset \mathcal{U}_k$ av dimension $k = m+1$ och $l = m$ som svarar mot grund- respektive hypotesmodellerna. Därför kommer även $\hat{\hat{\boldsymbol{\mu}}}$ vara projektionen av $\hat{\boldsymbol{\mu}}$ ned på hypotesrummet $\mathcal{U}_l$. Eftersom hypotesmodellen innehåller intercept så gäller $\bar{Y} \in \mathcal{U}_l$ och därmed också $\hat{\hat{\boldsymbol{\mu}}} - \bar{Y} \in \mathcal{U}_l$. Eftersom $\hat{\boldsymbol{\mu}} - \hat{\hat{\boldsymbol{\mu}}}$ är ortogonal mot alla element i $\mathcal{U}_l$ så är $\hat{\boldsymbol{\mu}} - \hat{\hat{\boldsymbol{\mu}}}$ ortogonal mot $\hat{\hat{\boldsymbol{\mu}}} - \bar{Y}$. Av detta följer att

$$\|\hat{\boldsymbol{\mu}} - \bar{Y}\|^2 = \|(\hat{\boldsymbol{\mu}} - \hat{\hat{\boldsymbol{\mu}}}) + (\hat{\hat{\boldsymbol{\mu}}} - \bar{Y})\|^2 = \|\hat{\boldsymbol{\mu}} - \hat{\hat{\boldsymbol{\mu}}}\|^2 + \|\hat{\hat{\boldsymbol{\mu}}} - \bar{Y}\|^2.$$

Genom insättning i (4)-(5) ger det i sin tur att

$$R_0^2 - R_1^2 = \frac{\|\hat{\boldsymbol{\mu}} - \hat{\hat{\boldsymbol{\mu}}}\|^2}{\|\boldsymbol{Y} - \bar{Y}\|^2} = \frac{\sum_{i=1}^N (\hat{\mu}_i - \hat{\hat{\mu}}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}. \tag{6}$$

c) Eftersom minsta kvadrat-skattningen av $\boldsymbol{\theta}$ ges av $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta}_1, \ldots, \hat{\beta}_m)^T = (\bar{Y}, \hat{\beta}_1, \ldots, \hat{\beta}_m)^T$, där $\hat{\beta}_j$ svarar mot kolumnen $\boldsymbol{x}_j - \bar{\boldsymbol{x}}_j$ i designmatrisen $\boldsymbol{A}$, så följer att

$$\hat{\boldsymbol{\mu}} = \boldsymbol{A}\hat{\boldsymbol{\theta}} = \bar{Y} + \sum_{l=1}^m \hat{\beta}_l (\boldsymbol{x}_l - \bar{\boldsymbol{x}}_l) = \hat{\beta}_j \boldsymbol{x}_j + \boldsymbol{v}, \tag{7}$$

där $\boldsymbol{v} \in \mathcal{U}_l$. Det beror på att $\bar{\boldsymbol{x}}_j$ och alla kolumner i $\boldsymbol{A}$ som inte svarar mot kovariat j, tillhör hypotesrummet $\mathcal{U}_l$. Nu är $\hat{\hat{\boldsymbol{x}}}_j$ en projektion av $\boldsymbol{x}_j$ ned på hypotesrummet $\mathcal{U}_l$. Således är $\boldsymbol{x}_j - \hat{\hat{\boldsymbol{x}}}_j$, och därmed även $\hat{\beta}_j(\boldsymbol{x}_j - \hat{\hat{\boldsymbol{x}}}_j)$, ortogonal mot $\mathcal{U}_l$. Vi skriver om (7) som

$$\hat{\boldsymbol{\mu}} = \hat{\beta}_j(\boldsymbol{x}_j - \hat{\hat{\boldsymbol{x}}}_j) + \boldsymbol{w}, \tag{8}$$

där $\boldsymbol{w} = \boldsymbol{v} - \hat{\beta}_j \hat{\hat{\boldsymbol{x}}}_j \in \mathcal{U}_l$, eftersom $\boldsymbol{v} \in \mathcal{U}_l$ och $\hat{\hat{\boldsymbol{x}}}_j \in \mathcal{U}_l$. Men eftersom projektionen av $\hat{\boldsymbol{\mu}}$ ned på hypotesrummet är $\hat{\hat{\boldsymbol{\mu}}}$, och $\hat{\beta}_j(\boldsymbol{x}_j - \hat{\hat{\boldsymbol{x}}}_j)$ är ortogonal mot $\mathcal{U}_l$, så följer att $\boldsymbol{w} = \hat{\hat{\boldsymbol{\mu}}}$ i (8). Därmed har vi visat att

$$\hat{\mu}_i = \hat{\beta}_j(x_{ji} - \hat{\hat{x}}_{ji}) + \hat{\hat{\mu}}_i$$

för $i = 1, \ldots, N$. Insättning i (6) ger

$$R_0^2 - R_1^2 = \frac{\hat{\beta}_j^2 \sum_{i=1}^N (x_{ji} - \hat{\hat{x}}_{ji})^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}.$$

Hur mycket mer vi förklarar med hjälp av $\boldsymbol{x}_j$ beror alltså dels på hur stor skattad effekt $\hat{\beta}_j$ denna variabel har och dels på hur stor del av kovariatvektorn $\boldsymbol{x}_j$ som inte förklaras av de övriga kovariaterna $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_{j+1}, \ldots, \boldsymbol{x}_m$.