STAT 510 Homework 13

Ungraded

1. (a) Specify matrix $\boldsymbol{X}$:

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{n_1\times 1}\otimes \boldsymbol{I}_t & & \\ & \mathbf{1}_{n_2\times 1}\otimes \boldsymbol{I}_t & \\ & & \mathbf{1}_{n_3\times 1}\otimes \boldsymbol{I}_t \end{bmatrix}$$

(b) Specify matrix $\text{Var}(\boldsymbol{y}) = \boldsymbol{\Sigma}$ in terms of $\boldsymbol{W}$:

$$\boldsymbol{\Sigma} = \boldsymbol{I}_{(n_1+n_2+n_3)}\otimes \boldsymbol{W}$$

(c) Compute $(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}$:

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{I}_{(n_1+n_2+n_3)}\otimes \boldsymbol{W}^{-1}$$

$$\boldsymbol{X}' = \begin{bmatrix} \mathbf{1}_{1\times n_1}\otimes \boldsymbol{I}_t & & \\ & \mathbf{1}_{1\times n_2}\otimes \boldsymbol{I}_t & \\ & & \mathbf{1}_{1\times n_3}\otimes \boldsymbol{I}_t \end{bmatrix}$$

$$\begin{aligned}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} &= \begin{bmatrix} (\mathbf{1}_{1\times n_1}\cdot \boldsymbol{I}_{n_1}\cdot \mathbf{1}_{n_1\times 1})\otimes(\boldsymbol{I}_t\cdot \boldsymbol{W}^{-1}\cdot \boldsymbol{I}_t) & & \\ & (\mathbf{1}_{1\times n_2}\cdot \boldsymbol{I}_{n_2}\cdot \mathbf{1}_{n_2\times 1})\otimes(\boldsymbol{I}_t\cdot \boldsymbol{W}^{-1}\cdot \boldsymbol{I}_t) & \\ & & (\mathbf{1}_{1\times n_3}\cdot \boldsymbol{I}_{n_3}\cdot \mathbf{1}_{n_3\times 1})\otimes(\boldsymbol{I}_t\cdot \boldsymbol{W}^{-1}\cdot \boldsymbol{I}_t) \end{bmatrix}\\ &= \begin{bmatrix} n_1\boldsymbol{W}^{-1} & & \\ & n_2\boldsymbol{W}^{-1} & \\ & & n_3\boldsymbol{W}^{-1} \end{bmatrix}\end{aligned}$$

therefore

$$(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} = \begin{bmatrix} \frac{\boldsymbol{W}}{n_1} & & \\ & \frac{\boldsymbol{W}}{n_2} & \\ & & \frac{\boldsymbol{W}}{n_3} \end{bmatrix}$$

(d) Compute $(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}$:

$$\begin{aligned}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1} &= \begin{bmatrix} (\mathbf{1}_{1\times n_1}\cdot \boldsymbol{I}_{n_1})\otimes(\boldsymbol{I}_t\cdot \boldsymbol{W}^{-1}) & & \\ & (\mathbf{1}_{1\times n_2}\cdot \boldsymbol{I}_{n_2})\otimes(\boldsymbol{I}_t\cdot \boldsymbol{W}^{-1}) & \\ & & (\mathbf{1}_{1\times n_3}\cdot \boldsymbol{I}_{n_3})\otimes(\boldsymbol{I}_t\cdot \boldsymbol{W}^{-1}) \end{bmatrix}\\ &= \begin{bmatrix} \mathbf{1}_{1\times n_1}\otimes \boldsymbol{W}^{-1} & & \\ & \mathbf{1}_{1\times n_2}\otimes \boldsymbol{W}^{-1} & \\ & & \mathbf{1}_{1\times n_3}\otimes \boldsymbol{W}^{-1} \end{bmatrix}\end{aligned}$$

so

$$(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \frac{1}{n_1}\mathbf{1}_{1\times n_1}\otimes \boldsymbol{I}_t & & \\ & \frac{1}{n_2}\mathbf{1}_{1\times n_2}\otimes \boldsymbol{I}_t & \\ & & \frac{1}{n_3}\mathbf{1}_{1\times n_3}\otimes \boldsymbol{I}_t \end{bmatrix}$$

(e) Compute $(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$:

$$
\begin{aligned}
(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} &= \begin{bmatrix} \frac{1}{n_1}\mathbf{1}_{1\times n_1}\otimes \boldsymbol{I}_t & & \\ & \frac{1}{n_2}\mathbf{1}_{1\times n_2}\otimes \boldsymbol{I}_t & \\ & & \frac{1}{n_3}\mathbf{1}_{1\times n_3}\otimes \boldsymbol{I}_t \end{bmatrix} \cdot \boldsymbol{y} \\
&= \begin{bmatrix} \frac{1}{n_1}\sum_{j=1}^{n_1}\boldsymbol{y}_{1j} \\ \frac{1}{n_2}\sum_{j=1}^{n_2}\boldsymbol{y}_{2j} \\ \frac{1}{n_3}\sum_{j=1}^{n_3}\boldsymbol{y}_{3j} \end{bmatrix}
\end{aligned}
$$

(f) Give the BLUEs of $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3$:

$$
\begin{aligned}
\hat{\boldsymbol{\mu}}_1 &= (\boldsymbol{I}_t, \mathbf{0}_{t\times t}, \mathbf{0}_{t\times t})\hat{\boldsymbol{\beta}}_{\text{OLS}} \\
&= (\boldsymbol{I}_t, \mathbf{0}_{t\times t}, \mathbf{0}_{t\times t})(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \\
&= \frac{1}{n_1}\sum_{j=1}^{n_1}\boldsymbol{y}_{1j}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\hat{\boldsymbol{\mu}}_2 &= (\mathbf{0}_{t\times t}, \boldsymbol{I}_t, \mathbf{0}_{t\times t})\hat{\boldsymbol{\beta}}_{\text{OLS}} \\
&= (\mathbf{0}_{t\times t}, \boldsymbol{I}_t, \mathbf{0}_{t\times t})(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \\
&= \frac{1}{n_2}\sum_{j=1}^{n_2}\boldsymbol{y}_{2j} \\
\hat{\boldsymbol{\mu}}_3 &= (\mathbf{0}_{t\times t}, \mathbf{0}_{t\times t}, \boldsymbol{I}_t)\hat{\boldsymbol{\beta}}_{\text{OLS}} \\
&= (\mathbf{0}_{t\times t}, \mathbf{0}_{t\times t}, \boldsymbol{I}_t)(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} \\
&= \frac{1}{n_3}\sum_{j=1}^{n_3}\boldsymbol{y}_{3j}
\end{aligned}
$$

2.

```
> y=c(15,9,15,23,14,18,5,7,12,11)
>
> o=glm(y~1,family=poisson(link=log))
>
> 1-pchisq(deviance(o),df.residual(o))
[1] 0.01685265
>
> #The residual deviance statistic suggests
> #that there is significant lack of fit.
> #The p-value is 0.01685.
>
> #The Pearson statistic is
>
> P=sum((y-mean(y))^2/mean(y))
> P
[1] 19.75969
>
```

```
> 2*(1-pchisq(P,length(y)-1))
[1] 0.03890952
>
> #The Pearson statistic also suggests
> #that there is significant lack of fit.
>
> #We should conclude that the data are not
> #an independent and identically distributed
> #sample from one Poisson distribution.
```

3.

```
> y=c(39, 31, 43, 31, 34, 36, 34, 24,
+ 23, 28, 24, 19, 16, 20, 25, 12,
+ 36, 38, 33, 22, 23, 17, 29, 16)
>
> g=as.factor(rep(c("A","B","C"),each=8))
>
> o=glm(y~g,family=poisson(link=log))
>
> o

Call:  glm(formula = y ~ g, family = poisson(link = log))

Coefficients:
(Intercept)           gB           gC
     3.5264      -0.4878      -0.2398

Degrees of Freedom: 23 Total (i.e. Null);  21 Residual
Null Deviance:      61.02
Residual Deviance: 35.57        AIC: 163.9
>
> anova(o,test="Chisq")
Analysis of Deviance Table

Model: poisson, link: log

Response: y

Terms added sequentially (first to last)


     Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                    23     61.017
g     2   25.452        21     35.565 2.973e-06 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
>
> #The test above suggests that there are
```

```
> #significant differences among genotypes.
>
> #Before going further with analysis,
> #let's check for overdispersion.
>
> 1-pchisq(deviance(o),df.residual(o))
[1] 0.02445231
>
> #The test suggests a lack of fit that
> #could be caused by over dispersion.
>
> #Let's look at a residual plot to make sure
> #the lack of fit is not due to extreme outliers.
>
> plot(fitted(o),resid(o,type="deviance"))
>
> #No extreme outliers noted. Thus, it seems
> #reasonable to blame the lack of fit on
> #overdispersion.
>
> #Let's estimate overdispersion parameter.
>
> phihat=deviance(o)/df.residual(o)
> phihat
[1] 1.693594
>
> #Let's test again for a difference among
> #genotypes, but this time we will account
> #for overdispersion
>
> oq=glm(y~g,family=quasipoisson(link=log))
>
> anova(oq,test="F")
Analysis of Deviance Table

Model: quasipoisson, link: log

Response: y

Terms added sequentially (first to last)


     Df Deviance Resid. Df Resid. Dev      F   Pr(>F)
NULL                    23     61.017
g     2   25.452        21     35.565 7.7292 0.003051 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

```
>
> #There is significant evidence of differences
> #among genotypes.
>
> #Let's compare pairs of genotypes.
>
> v=vcov(oq)
> b=coef(oq)
>
> C=matrix(c(
+ 0,1,0,
+ 0,0,1,
+ 0,1,-1),byrow=T,nrow=3)
>
> Cb=C%*%b
> se=sqrt(diag(C%*%v%*%t(C)))
> tt=drop(Cb/se)
> 2*(1-pt(abs(tt),df.residual(o)))
[1] 0.0008923671 0.0535405334 0.0752392327
>
> #Based on the p-values above, all pairwise
> #comparisons are significant at the .10 level.
> #Only A vs. B is significant at the .05 level.
>
> coef(oq)
(Intercept)          gB          gC
  3.5263605  -0.4878083  -0.2398261
>
> #Genotype A seems significantly more susceptible
> #then genotype B.
>
> #Now let's address overdispersion by fitting a
> #GLMM that allows for overdispersion in the data.
>
> library(lme4)
Loading required package: lattice
Loading required package: Matrix
Warning message:
package lme4 was built under R version 2.15.3
>
> leaf=factor(1:24)
> oglmm=glmer(y~g+(1|leaf),family=poisson(link="log"))
> oglmmreduced=glmer(y~1+(1|leaf),family=poisson(link="log"))
> anova(oglmmreduced,oglmm)
Data:
Models:
oglmmreduced: y ~ 1 + (1 | leaf)
```

```
oglmm: y ~ g + (1 | leaf)
             Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
oglmmreduced  2 173.52 175.88 -84.761   169.52
oglmm         4 164.26 168.97 -78.129   156.26 13.264      2   0.001318 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
>
> #From the above, we see that the likelihood ratio test
> #statistic for comparing the null model with only an
> #intercept parameter and a leaf variance component
> #to the alternative model with one parameter for each
> #genotype and a leaf variance component is 13.264.
> #Comparing to a chi-square distribution with 2 df
> #results in a p-value of 0.001318.
```

4.

```
> library(lme4)
>
> mu1 = 3
> mu2 = 3
> n1 = 5
> n2 = 5
> sigma = .25
> N = 10000
> obs = factor(1 : (n1 + n2))
> trt = factor(rep(1:2, c(n1, n2)))
> trt
 [1] 1 1 1 1 1 2 2 2 2 2
Levels: 1 2
> stat1 = rep(0, N)
> stat2 = rep(0, N)
> stat3 = rep(0, N)
>
> set.seed(82361)
> for(i in 1:N){
+   lambda1 = exp(mu1 + rnorm(n1, 0, sigma))
+   lambda2 = exp(mu2 + rnorm(n2, 0, sigma))
+   y = c(rpois(n1, lambda1), rpois(n2, lambda2))
+   oglm = glm(y ~ trt, family = poisson(link = "log"))
+   phihat = deviance(oglm) / df.residual(oglm)
+   b = coef(oglm)
+   se = sqrt(vcov(oglm)[2,2])
+   stat1[i] = b[2] / se
+   stat2[i] = b[2] / (sqrt(phihat) * se)
+   oglmer = glmer(y ~ trt + (1 | obs), family = poisson(link = "log"))
+   stat3[i] = fixef(oglmer)[2] / sqrt(vcov(oglmer)[2,2])
+ }
Warning messages:
```

```
1: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio
 - Rescale variables?
2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio
 - Rescale variables?
3: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio
 - Rescale variables?
4: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio
 - Rescale variables?
5: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio
 - Rescale variables?
6: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,  :
  Model is nearly unidentifiable: large eigenvalue ratio
 - Rescale variables?
>
> p1 = 2 * (1 - pnorm(abs(stat1), 0, 1))
> p2 = 2 * (1 - pt(abs(stat2), (n1 + n2 - 2)))
> p3 = 2 * (1 - pnorm(abs(stat3), 0, 1))
>
> mean(p1 <= 0.05)
[1] 0.2021
> mean(p2 <= 0.05)
[1] 0.0496
> mean(p3 <= 0.05)
[1] 0.105
```

The above code and output indicates that the quasilikelihood approach (Test 2 in the problem statement) is the only one of the three approaches that controls the type I error rate at the nominal 0.05 level. The Test 1 approach (GLM ignoring overdispersion) rejects a true null hypothesis around 20% of the time rather than the 5% that should occur when using p-value 0.05 as the threshold for significance. The GLMM approach (Test 3) is better than the GLM ignoring overdispersion, but the type I error rate (approximately 10.5%) is still twice what it should be. The GLMM approach has some numerical convergence problems for a few of the 10,000 simulated datasets (which is the cause of the warning messages), but these problems do not affect the general conclusion that the GLMM does not control the type I error rate at the 5% level. We should expect the GLMM (Test 3) to improve as the sample sizes (n_1 and n_2) increase. However, the GLM (Test 1) will not likely improve with increasing sample size because the model is wrong about the variance of the responses. As σ (which controls the extent of overdispersion) decreases towards 0 and the sample sizes grow, the GLM is expected to perform better.

5.
```
> d=read.delim(
+ "http://www.public.iastate.edu/~dnett/S510/PlaneCrashes.txt")
> d
```

```
   index crashes
1    376       8
2    347       5
3    322       8
4    104       4
5    103       6
6     98       4
7     96       8
8     85       6
9     82       4
10    63       2
11    44       7
12    40       4
13     5       3
14     5       2
15     0       4
16     0       3
17     0       2
>
> plot(d)
>
> o=glm(crashes~index,family=poisson(link=log),data=d)
>
> summary(o)

Call:
glm(formula = crashes ~ index, family = poisson(link = log),
    data = d)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1974  -0.3978  -0.1766   0.3537   1.4919

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) 1.3098588  0.1582327   8.278   <2e-16 ***
index       0.0019933  0.0008166   2.441   0.0146 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 15.295  on 16  degrees of freedom
Residual deviance:  9.794  on 15  degrees of freedom
AIC: 70.365

Number of Fisher Scoring iterations: 4
```

```
>
> anova(o,test="Chisq")
Analysis of Deviance Table

Model: poisson, link: log

Response: crashes

Terms added sequentially (first to last)


      Df Deviance Resid. Df Resid. Dev Pr(>Chi)
NULL                     16     15.295
index  1   5.5013        15      9.794    0.019 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
>
> #It looks like there is significant evidence
> #of association between the news coverage index
> #and the number of crashes. This might be evidence
> #in favor of these sociologists’ theory.
>
> #Check for lack of fit.
>
> 1-pchisq(deviance(o),df.residual(o))
[1] 0.8324938
>
> #There is no evidence of lack of fit.
> #However, it’s not clear how good the
> #asymptotic chi-square approximation
> #will be in this case since n is low
> #and the counts are small.
>
> exp(100*coef(o)[2])
  index
1.22059
>
> #A 100 unit increase in news coverage index
> #is associated with an estimated 22% increase
> #in the mean number of crashes that occur in the
> #subsequent week.
```

6. (a) The model in this problem is $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ with $\boldsymbol{\beta} = (\mu, \alpha_1, \alpha_2, \beta_1, \beta_2)'$ and

$$\boldsymbol{X} = \left(\mathbf{1}_{20\times1}, \begin{bmatrix} \mathbf{1}_{10\times1} \\ \mathbf{0}_{10\times1} \end{bmatrix}, \begin{bmatrix} \mathbf{0}_{10\times1} \\ \mathbf{1}_{10\times1} \end{bmatrix}, \begin{bmatrix} \mathbf{1}_{2\times1} \\ \mathbf{0}_{8\times1} \\ \mathbf{1}_{6\times1} \\ \mathbf{0}_{4\times1} \end{bmatrix}, \begin{bmatrix} \mathbf{0}_{2\times1} \\ \mathbf{1}_{8\times1} \\ \mathbf{0}_{6\times1} \\ \mathbf{1}_{4\times1} \end{bmatrix} \right)$$

Because $rank(\boldsymbol{X}) = 3$, one possible full rank matrix is

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \left(\mathbf{1}_{20\times1}, \begin{bmatrix} \mathbf{1}_{10\times1} \\ \mathbf{0}_{10\times1} \end{bmatrix}, \begin{bmatrix} \mathbf{1}_{2\times1} \\ \mathbf{0}_{8\times1} \\ \mathbf{1}_{6\times1} \\ \mathbf{0}_{4\times1} \end{bmatrix} \right)$$

We can compute orthogonal columns using Gram-Schmidt Orthogonalization method:

$$\begin{aligned} \boldsymbol{w}_1 &= \boldsymbol{x}_1; \\ \boldsymbol{w}_2 &= (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{w}_1})\boldsymbol{x}_2 = \frac{1}{2} \begin{bmatrix} \mathbf{1}_{10\times1} \\ -\mathbf{1}_{10\times1} \end{bmatrix} \\ \boldsymbol{w}_3 &= (\boldsymbol{I} - \boldsymbol{P}_{[\boldsymbol{w}_1, \boldsymbol{w}_2]})\boldsymbol{x}_3 = \begin{bmatrix} 0.8 \cdot \mathbf{1}_{2\times1} \\ -0.2 \cdot \mathbf{1}_{8\times1} \\ 0.4 \cdot \mathbf{1}_{6\times1} \\ -0.6 \cdot \mathbf{1}_{4\times1} \end{bmatrix} \end{aligned}$$

So a model matrix $\boldsymbol{W}$ with orthogonal columns is $\boldsymbol{W} = (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$.

(b) Compute $\boldsymbol{P_W y}$:

$$\begin{aligned} (\boldsymbol{W}'\boldsymbol{W})^{-1} &= \left[\begin{pmatrix} \boldsymbol{w}_1' \\ \boldsymbol{w}_2' \\ \boldsymbol{w}_3' \end{pmatrix} (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3) \right]^{-1} = \begin{bmatrix} \boldsymbol{w}_1'\boldsymbol{w}_1 & & \\ & \boldsymbol{w}_2'\boldsymbol{w}_2 & \\ & & \boldsymbol{w}_3'\boldsymbol{w}_3 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 20 & & \\ & 5 & \\ & & 4 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{20} & & \\ & \frac{1}{5} & \\ & & \frac{1}{4} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \boldsymbol{W}'\boldsymbol{y} &= \begin{pmatrix} \boldsymbol{w}_1' \\ \boldsymbol{w}_2' \\ \boldsymbol{w}_3' \end{pmatrix} \cdot \boldsymbol{y} = \begin{bmatrix} \boldsymbol{w}_1'\boldsymbol{y} \\ \boldsymbol{w}_2'\boldsymbol{y} \\ \boldsymbol{w}_3'\boldsymbol{y} \end{bmatrix} \\ &= \begin{bmatrix} \sum_i \sum_j \sum_k y_{ijk} \\ \frac{1}{2}(\sum_j \sum_k y_{1jk} - \sum_j \sum_k y_{2jk}) \\ 0.8 \cdot \sum_k y_{11k} - 0.2 \cdot \sum_k y_{12k} + 0.4 \cdot \sum_k y_{21k} - 0.6 \cdot \sum_k y_{22k} \end{bmatrix} \\ &= \begin{bmatrix} 100 \\ -4 \\ 6.4 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
\boldsymbol{P_W y} &= \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{y} \\
&= (\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3) \begin{bmatrix} \frac{1}{20} & & \\ & \frac{1}{5} & \\ & & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 100 \\ -4 \\ 6.4 \end{bmatrix} \\
&= 5\boldsymbol{w}_1 - 0.8\boldsymbol{w}_2 + 1.6\boldsymbol{w}_3 \\
&= \begin{bmatrix} 5.88 \cdot \mathbf{1}_{2\times 1} \\ 4.28 \cdot \mathbf{1}_{8\times 1} \\ 6.04 \cdot \mathbf{1}_{6\times 1} \\ 4.44 \cdot \mathbf{1}_{4\times 1} \end{bmatrix}
\end{aligned}$$

(c) The Type II sum of squares for factor B is $S(B|1, A) = \boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}$ where $\boldsymbol{P}_3\boldsymbol{y} = \boldsymbol{P_W y}$ and from part (b)

$$\begin{aligned}
\boldsymbol{P}_2\boldsymbol{y} &= \boldsymbol{P}_{[\boldsymbol{w}_1, \boldsymbol{w}_2]}\boldsymbol{y} \\
&= (\boldsymbol{w}_1, \boldsymbol{w}_2) \begin{bmatrix} \frac{1}{20} & \\ & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 100 \\ -4 \end{bmatrix} \\
&= 5\boldsymbol{w}_1 - 0.8\boldsymbol{w}_2 \\
&= \begin{bmatrix} 4.6 \cdot \mathbf{1}_{10\times 1} \\ 5.4 \cdot \mathbf{1}_{10\times 1} \end{bmatrix}
\end{aligned}$$

so

$$\begin{aligned}
S(B|1, A) &= \boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}_3\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{P}_2\boldsymbol{y} = ||\boldsymbol{P}_3\boldsymbol{y}||^2 - ||\boldsymbol{P}_2\boldsymbol{y}||^2 \\
&= (5.88^2 \times 2 + 4.28^2 \times 8 + 6.04^2 \times 6 + 4.44^2 \times 4) - (4.6^2 \times 10 + 5.4^2 \times 10) \\
&= 10.24
\end{aligned}$$