MATH-542 Final Exam

Saket Choudhary

4/30/2016

```r
chemical.data <- read.csv('chemical.csv', header=T)
N <- nrow(chemical.data)
model <- lm(y1 ~ x1 + x2 + x3, data=chemical.data)
model.summary <- summary(model)
print(xtable(model.summary, caption="Model 1 Summary"))
```

% latex table generated in R 3.2.2 by xtable 1.8-2 package %

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	332.1110	18.6929	17.77	0.0000
x1	-1.5460	0.0990	-15.61	0.0000
x2	-1.4246	0.1476	-9.65	0.0000
x3	-2.2374	0.3398	-6.58	0.0000

Part (2.1)

```r
beta <- coef(model)
sigma <- summary(model)$sigma
sigma.squared <- sigma**2
```

β: (Intercept x1 x2 x3):

```
##
##
## ------------  -----------
## (Intercept)    332.110983
## x1              -1.545961
## x2              -1.424559
## x3              -2.237366
## ------------  -----------
```

σ^2:

5.3449026

Part (2.2)

Covariance matrix of the coefficients:

```r
cov.beta <- vcov(model)
print(xtable(cov.beta, caption="Part 2.2 Covariance matrix of coefficients"))
```

% latex table generated in R 3.2.2 by xtable 1.8-2 package %

	(Intercept)	x1	x2	x3
(Intercept)	349.43	-1.81	-1.67	-0.11
x1	-1.81	0.01	0.01	-0.00
x2	-1.67	0.01	0.02	-0.01
x3	-0.11	-0.00	-0.01	0.12

Part (2.3)

```
R.unadjusted <- model.summary$r.squared
R.adjusted <- model.summary$adj.r.squared
```

Coefficient of determination:

R^2 0.9551434 $R^2 adjusted$: 0.9461721

Part (2.4)

```
model2 <- lm(y1 ~ x1+x2+x3+I(x1^2)+I(x2^2)+I(x3^2)+x1*x2+x1*x3+x2*x3, data=chemical.data)
model2.summary <- summary(model2)
print(xtable(model2.summary, caption= "Model 2 Summary"))
```

% latex table generated in R 3.2.2 by xtable 1.8-2 package %

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	964.9291	645.4219	1.50	0.1691
x1	-7.4421	7.0832	-1.05	0.3208
x2	-11.5077	8.0830	-1.42	0.1883
x3	-2.1401	15.3457	-0.14	0.8922
I(x1^2)	0.0125	0.0200	0.62	0.5479
I(x2^2)	0.0332	0.0507	0.66	0.5284
I(x3^2)	-0.2940	0.2248	-1.31	0.2233
x1:x2	0.0535	0.0359	1.49	0.1707
x1:x3	0.0380	0.0992	0.38	0.7103
x2:x3	-0.1016	0.1512	-0.67	0.5183

```
beta2 <- coef(model2)
beta2.constant <- beta2[1]-3
beta2.adj.constant <- c(beta2.constant, beta2[-1])
sigma2 <- summary(model2)$sigma
sigma2.squared <- sigma2**2
```

Thus, $\hat{\beta}$: and σ^2:

Part (2.5)

```
R2.unadjusted <- model2.summary$r.squared
R2.adjusted <- model2.summary$adj.r.squared
```

For the second model:

R^2 0.9741468 $R^2 adjusted$: 0.9482936

Part (2.6)

```
print(xtable(anova(model2, model, test='F')))
```

% latex table generated in R 3.2.2 by xtable 1.8-2 package %

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	9	46.21				
2	15	80.17	-6	-33.97	1.10	0.4295

Part (2.7)

```
CI <- confint(model2, level=0.95, c(2,3,4))
CI
```

```
##         2.5 %    97.5 %
## x1 -23.46535  8.581090
## x2 -29.79275  6.777342
## x3 -36.85452 32.574268
```

Check this:

```
p <- 10
k <- 3
alpha <- 0.05
coef <- model2.summary$coefficients[,1][c(2,3,4)]
err <- model2.summary$coefficients[,2][c(2,3,4)]
coef - err * qt(1-alpha/2, N-p) ## same as CI above
```

```
##        x1        x2        x3
## -23.46535 -29.79275 -36.85452
```

Thus,95% CI for $\beta_0, \beta_1, \beta_2, \beta_3$:

Part (2.8)

```
coef <- model2.summary$coefficients[,1][c(2,3,4)]
std.err.coef <- model2.summary$coefficients[,2][c(2,3,4)]
t.stats <- qt(1-alpha/(2*N),df=N-p)
```

```r
CI.bonferroni.ub <-  coef + std.err.coef * t.stats
CI.bonferroni.lb <-  coef - std.err.coef * t.stats
CI.bonferroni <- c(CI.bonferroni.lb, CI.bonferroni.ub)
CI.bonferroni
```

```
##        x1        x2        x3        x1        x2        x3
## -36.56369 -44.74005 -65.23217  21.67944  21.72464  60.95191
```

95% bonferroni CI for $\beta_0, \beta_1, \beta_2, \beta_3$:

Part (2.9)

```r
range.data <- max(chemical.data$y1)-min(chemical.data$y1)
tukey.stats <- qtukey(1-alpha/2, nmeans=k  ,df=N-p)
CI.tukeys.ub <- coef + std.err.coef * tukey.stats
CI.tukeys.lb <- coef - std.err.coef * tukey.stats
CI.tukeys <- c(CI.tukeys.lb, CI.tukeys.ub)
CI.tukeys
```

```
##        x1        x2        x3        x1        x2        x3
## -39.87027 -48.51338 -72.39589  24.98601  25.49798  68.11563
```

Tukey's 95% CI :

Part (2.10)

```r
kfk <- k*qf(1-alpha/2, df1=k, df2=p)
CI.scheffes.ub <- coef + std.err.coef * (kfk**0.5)
CI.scheffes.lb <- coef - std.err.coef * (kfk**0.5)
CI.scheffes <- c(CI.scheffes.lb, CI.scheffes.ub)
CI.scheffes
```

```
##        x1        x2        x3        x1        x2        x3
## -34.39247 -42.26234 -60.52819  19.50821  19.24693  56.24794
```

Scheffe's's 95% CI :

Part (2.11)

```r
xob <- c(1,165,32,5) %*% beta2[1:4]
xob.CI.ub <- xob + t.stats * sigma
xob.CI.lb <- xob - t.stats * sigma
xob.CI <- c(xob.CI.lb, xob.CI.ub)
print (xob.CI)
```

```
## [1] -651.4743 -632.4641
```

95% CI:

Part (2.12)

```
xob.PI.ub <- xob + t.stats * sigma * sqrt(1+1/N)
xob.PI.lb <- xob - t.stats * sigma * sqrt(1+1/N)
xob.PI <- c(xob.PI.lb, xob.PI.ub)
print (xob.PI)
```

```
## [1] -651.7212 -632.2171
```

96% Prediction Interval:

Part (2.13)

```
leverage <- hat(model.matrix(model2))
epsilon <- residuals(model2)
prediction <- predict(model2)
studentised.residual <- rstandard(model2)
cooksdistance <- cooks.distance(model2)
deleted.studentised.residual <- rstudent(model2)
df <- data.frame(y_i=chemical.data$y1,
                 y_hat= prediction,
                 epsilon=epsilon,
                 h=leverage,
                 r=studentised.residual,
                 t=studentised.residual,
                 D=cooksdistance)
```

```
colnames(df) <- c('$y_i$', '$\\hat{y}_i$',
                  '$\\epsilon_i$', '$h_{ii}$',
                  '$r_i$' , '$t_i$', '$D_i$')
html <- xtable(df, auto=TRUE)
print(html, sanitize.text.function=function(x){x},
      size = "\\setlength{\\tabcolsep}{12pt}")
```

% latex table generated in R 3.2.2 by xtable 1.8-2 package %

	y_i	$\hat{y}_i$	ϵ_i	h_{ii}	r_i	t_i	D_i
1	41.5	40.8997843	0.6002157	0.9264479	0.9767224	0.9767224	1.20162237
2	33.8	33.1529637	0.6470363	0.7680842	0.5929594	0.5929594	0.11644707
3	27.7	28.4551627	-0.7551627	0.5982579	-0.5258094	-0.5258094	0.04117160
4	21.7	19.9086921	1.7913079	0.3295035	0.9654581	0.9654581	0.04580684
5	19.9	18.5229104	1.3770896	0.5234560	0.8803840	0.8803840	0.08513761
6	15.0	12.6421354	2.3578646	0.3660300	1.3069108	1.3069108	0.09861429
7	12.2	13.5947505	-1.3947505	0.4926836	-0.8642085	-0.8642085	0.07253144
8	4.3	6.1894794	-1.8894794	0.4608305	-1.1356408	-1.1356408	0.11022952
9	19.3	20.6935489	-1.3935489	0.3519984	-0.7640040	-0.7640040	0.03170705
10	6.4	6.3117728	0.0882272	0.5994800	0.0615250	0.0615250	0.00056657
11	37.6	37.5667479	0.0332521	0.7144815	0.0274639	0.0274639	0.00018875
12	18.0	14.4959353	3.5040647	0.4180725	2.0272118	2.0272118	0.29524389
13	26.3	28.5521014	-2.2521014	0.4534991	-1.3444789	-1.3444789	0.15000080
14	9.9	10.8338004	-0.9338004	0.7038411	-0.7572741	-0.7572741	0.13628748
15	25.0	25.2610460	-0.2610460	0.5589153	-0.1734673	-0.1734673	0.00381293
16	14.1	14.7094284	-0.6094284	0.4585575	-0.3655174	-0.3655174	0.01131508
17	15.2	14.7094284	0.4905716	0.4585575	0.2942305	0.2942305	0.00733191
18	15.9	18.4501558	-2.5501558	0.4086516	-1.4635463	-1.4635463	0.14802082
19	19.6	18.4501558	1.1498442	0.4086516	0.6599009	0.6599009	0.03009313