1.(a) Since A is nonnegative definite (semipositive definite) matrix, we know by definition that

$$\underset{\sim}{y}^T A \underset{\sim}{y} \geq 0 \text{ for any vector } \underset{\sim}{y}.$$

For any vector $\underset{\sim}{z}$, let $\underset{\sim}{y} = P\underset{\sim}{z}$. Then,

$$\underset{\sim}{z}^T P^T A P \underset{\sim}{z} = (P\underset{\sim}{z})^T A\,(P\underset{\sim}{z}) = \underset{\sim}{y}^T A \underset{\sim}{y} \geq 0 \quad \text{for any vector} \quad \underset{\sim}{z}.$$

Consequently, $P^T AP$ satisfies the definition of a nonnegative definite matrix.
This result holds even when P is a singualr matrix.

(b) For this result, we will make use of the condition that P is a nonsingular matrix. Since A is a postive definite matrix, we have by definition that $\underset{\sim}{y}^T A \underset{\sim}{y} > 0$ for every $\underset{\sim}{y} \neq 0$. For any vector $\underset{\sim}{z}$, let $\underset{\sim}{y} = P\underset{\sim}{z}$. Then,

$$\underset{\sim}{z}^T P^T A P \underset{\sim}{z} = (P\underset{\sim}{z})^T A\,(P\underset{\sim}{z}) = \underset{\sim}{y}^T A \underset{\sim}{y} > 0 \quad \text{for every vector} \quad \underset{\sim}{y} = P\underset{\sim}{z} \neq 0.$$

Since P is nonsingular, the columns of P are linearly independent and $\underset{\sim}{y} = P\underset{\sim}{z} = 0$ only if $\underset{\sim}{z} = 0$. Consequently,$P^T AP$ satisfies the definition of a positive definite matrix.

2.(a) For B to satisfy the definition of a generalized inverse of A, we must have $ABA = A$. Although this simple example could be done by hand, to review some Splus we can do the following:

```
> b <- matrix(c(1,0,0,0),ncol=4)
> a <- matrix(c(1,2,5,-2),ncol=1)
> a %*% b %*% a
     [,1]
[1,]    1
[2,]    2
[3,]    5
[4,]   -2
```

(b) Two other generalized inverses for A are the following

$$G_1 = \left[0\ \tfrac{1}{2}\ 0\ 0\right], \quad G_2 = \left[0\ 0\ \tfrac{1}{5}\ 0\right]. \qquad \text{Check in each case that } AG_iA = A.$$

3.(a) and **(b)** By the definition of an eigenvector, $\underset{\sim}{x}$ is an eigenvector for $I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T$ if there is a scalar λ such that

$$(I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T)\underset{\sim}{x} \;=\; \lambda\underset{\sim}{x}$$

Note that

$$I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T)\underset{\sim}{x} \;=\; (\underset{\sim}{x} - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}(\underset{\sim}{x}^T\underset{\sim}{x}) \;=\; \underset{\sim}{x} - \underset{\sim}{x} \;=\; \underset{\sim}{0}$$

Consequently, $\underset{\sim}{x}$ is an eigenvector of $(I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T)$ corresponding to the eigenvalue $\lambda = 0$.

(c) Since $\underset{\sim}{x}^T\underset{\sim}{u} = 0$, we have $(I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T)\underset{\sim}{u} \;=\; \underset{\sim}{u}$.

Therefore, $\underset{\sim}{u}$ is an eigenvector of $(I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T)$ corresponding to an eigenvalue of $\lambda = 1$.

(d) From part (b) the eigenvalue associated with $\underset{\sim}{x}$ is 0. Furthermore, $\underset{\sim}{x}$ is the basis for a one dimensional space (a line). The space orthogonal to $\underset{\sim}{x}$ has dimension $n-1$ and we can find a set of vectors $\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_{n-1}$ that provide a basis for that space, are orthogonal to each other and are also orthogonal to $\underset{\sim}{x}$. Each of these vectors corresponds to and eigenvalue of 1. Hence, $I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T$ has one zero eigenvalue and the other $n-1$ eigenvalues are all 1. Furthermore,

$$rank(I - \underset{\sim}{x}(\underset{\sim}{x}^T\underset{\sim}{x})^{-1}\underset{\sim}{x}^T) \;=\; n-1,$$

the number of non-zero eigenvalues.

5.) Since A is an $n \times n$ symmetric matrix with $rank(A) = r$, we can use the spectral decomposition to write A as

$$\begin{aligned}
\underset{n\times n}{A} &= \underset{n\times n}{L} \begin{bmatrix} \Delta_r & \underset{r\times(n-r)}{\mathbf{0}} \\ \underset{(n-r)\times r}{\mathbf{0}} & \underset{(n-r)\times(n-r)}{\mathbf{0}} \end{bmatrix} \underset{n\times n}{L^T} \\
&= \begin{bmatrix} \underset{n\times r}{L_1} & \underset{p\times(n-r)}{L_2} \end{bmatrix} \begin{bmatrix} \Delta_r & \underset{r\times(n-r)}{\mathbf{0}} \\ \underset{(n-r)\times r}{\mathbf{0}} & \underset{(n-r)\times(n-r)}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} \underset{r\times n}{L_1^T} \\ \underset{(n-r)\times n}{L_2^T} \end{bmatrix} \\
&= \underset{n\times r}{L_1}\,\Delta_r\,\underset{r\times n}{L_1^T}
\end{aligned}$$

Note that

$$\begin{aligned}
\underset{n\times n}{G} &= \underset{n\times n}{L} \begin{bmatrix} \Delta_r^{-1} & \underset{r\times(n-r)}{\mathbf{0}} \\ \underset{(n-r)\times r}{\mathbf{0}} & \underset{(n-r)\times(n-r)}{\mathbf{0}} \end{bmatrix} \underset{n\times n}{L^T} \\
&= \begin{bmatrix} \underset{n\times r}{L_1} & \underset{n\times(n-r)}{L_2} \end{bmatrix} \begin{bmatrix} \Delta_r^{-1} & \underset{r\times(n-r)}{\mathbf{0}} \\ \underset{(n-r)\times r}{\mathbf{0}} & \underset{(n-r)\times(n-r)}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} \underset{r\times n}{L_1^T} \\ \underset{(n-r)\times n}{L_2^T} \end{bmatrix} \\
&= \underset{n\times r}{L_1}\,\Delta_r^{-1}\,\underset{r\times n}{L_1^T}
\end{aligned}$$

Now show that the four properties of the Moore-Penrose inverse are satisfied.

(i) $AGA = A$.

$$\begin{aligned}
AGA &= \left(\underset{n\times r}{L_1}\,\Delta_r\,\underset{r\times n}{L_1^T}\right)\left(\underset{n\times r}{L_1}\,\Delta_r^{-1}\,\underset{r\times n}{L_1^T}\right)\left(\underset{n\times r}{L_1}\,\Delta_r\,\underset{r\times n}{L_1^T}\right) \\
&= \underset{n\times r}{L_1}\,\Delta_r\left(\underset{r\times n}{L_1^T}\,\underset{n\times r}{L_1}\right)\Delta_r^{-1}\left(\underset{r\times n}{L_1^T}\,\underset{n\times r}{L_1}\right)\Delta_r\,\underset{r\times n}{L_1^T} \\
&= \underset{n\times r}{L_1}\,\Delta_r\Delta_r^{-1}\Delta_r\,\underset{r\times n}{L_1^T} \qquad \text{since } L_1^TL_1 = I_r \text{ and } L_1^TL_1 = I_r. \\
&= \underset{n\times r}{L_1}\,\Delta_r\,\underset{r\times n}{L_1^T} \qquad \text{since } \Delta_r\Delta_r^{-1} = I_r. \\
&= A
\end{aligned}$$

(ii) $GAG = G$. The proof is similar to (i).

(iii) AG is symmetric. i.e. $(AG)^T = AG$.

$$\begin{aligned} AG &= \underset{n\times r}{L_1}\,\Delta_r\,\underset{r\times n}{L_1^T}\;\underset{n\times r}{L_1}\,\Delta_r^{-1}\,\underset{r\times n}{L_1^T} \\ &= \underset{n\times r}{L_1}\,\Delta_r\,\underset{r\times n}{\left(L_1^T\,L_1\right)}\,\Delta_r^{-1}\,\underset{r\times n}{L_1^T} \\ &= \underset{n\times r}{L_1}\,\Delta_r\Delta_r^{-1}\,\underset{r\times n}{L_1^T} \quad since\ L_1^T L_1 = I_r. \\ &= \underset{n\times r}{L_1}\,\underset{r\times n}{L_1^T} \quad since\ \Delta_r\Delta_r^{-1} = I_r. \end{aligned}$$

Then (1)

$$(AG)^T = \left(\underset{n\times r}{L_1}\,\underset{r\times n}{L_1^T}\right)^T = \underset{r\times n}{(L_1^T)^T}\,\underset{n\times r}{(L_1)^T} = \underset{n\times r}{L_1}\,\underset{r\times n}{L_1^T} = AG\ .$$

(iv) GA is symmetric. i.e. $(GA)^T = GA$. The proof is similar to (iii).

5.(a)

```
> V <- matrix(c(3,-1,1,-1,5,-1,1,-1,3),ncol=3,byrow=T)
> V
     [,1] [,2] [,3]
[1,]    3   -1    1
[2,]   -1    5   -1
[3,]    1   -1    3
> eigen(V)
$values:
[1] 6 3 2

$vectors:
           [,1]      [,2]           [,3]
[1,]  0.4082483 0.5773503  7.071068e-001
[2,] -0.8164966 0.5773503  3.764350e-016
[3,]  0.4082483 0.5773503 -7.071068e-001
```

(b)

```
> ############################################################
> # Function: spectral
> #    input: V = Symmetric matrix
> #           p = Power(e.g. -1   -> inverse matrix
> #                          -1/2 -> inv. squart root matrix
> #   output: inverse or inverse square root matrix
> ############################################################
> spectral <- function(V,p)
+ {
+       eigen.V <- eigen(V)
+       eval    <- eigen.V$values
+       evec    <- eigen.V$vectors
+       spec.V  <- evec %*% diag(eval^p) %*% t(evec)
+       return(spec.V)
+ }
```

(c)

```
> VV <- spectral(V,-1/2)
> VV
            [,1]       [,2]        [,3]
[1,]  0.61404486 0.05636733 -0.09306192
[2,]  0.05636733 0.46461562  0.05636733
[3,] -0.09306192 0.05636733  0.61404486

#  The following should be an identity matrix
> VV %*% V %*% VV
```

```
              [,1]           [,2]           [,3]
[1,] 1.000000e+000 -1.943094e-018  8.010899e-017
[2,] 1.100126e-017  1.000000e+000 -3.498043e-016
[3,] 1.587679e-016 -3.041323e-016  1.000000e+000
```

(d) $V^{-1/2} = UD^{-1/2}U^T$, where $U = \begin{bmatrix} u_1 & u_2 & u_3 \\ \sim & \sim & \sim \end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix}$.

Then, for $\underset{\sim}{Z} = V^{-1/2}\underset{\sim}{Y}$ we have

$$\begin{aligned} \mathrm{E}(\underset{\sim}{Z}) &= \mathrm{E}(A^{-\frac{1}{2}}\underset{\sim}{Y}) = V^{-\frac{1}{2}}\mathrm{E}(\underset{\sim}{Y}) = V^{-\frac{1}{2}}\underset{\sim}{0} = \underset{\sim}{0} \\ \mathrm{Var}(\underset{\sim}{Z}) &= \mathrm{Var}(V^{-\frac{1}{2}}\underset{\sim}{Y}) = V^{-\frac{1}{2}}\mathrm{Var}(\underset{\sim}{Y})[V^{-\frac{1}{2}}]^T = V^{-\frac{1}{2}}VV^{-\frac{1}{2}} = I \end{aligned}$$

6.(a) Let $\underline{b} = (X^TX)^-X^T\underline{Y}$ for some generalized inverse $(X^TX)^-$ of X^TX. Then $X^TX\underline{b} = X^TX(X^TX)^-X^T\underline{Y} = X^T\underline{Y}$ because $X(X^TX)^-X^TX = P_XX = X$. (Some students based their argument on an incorrect claim that $P_XY = Y$.)

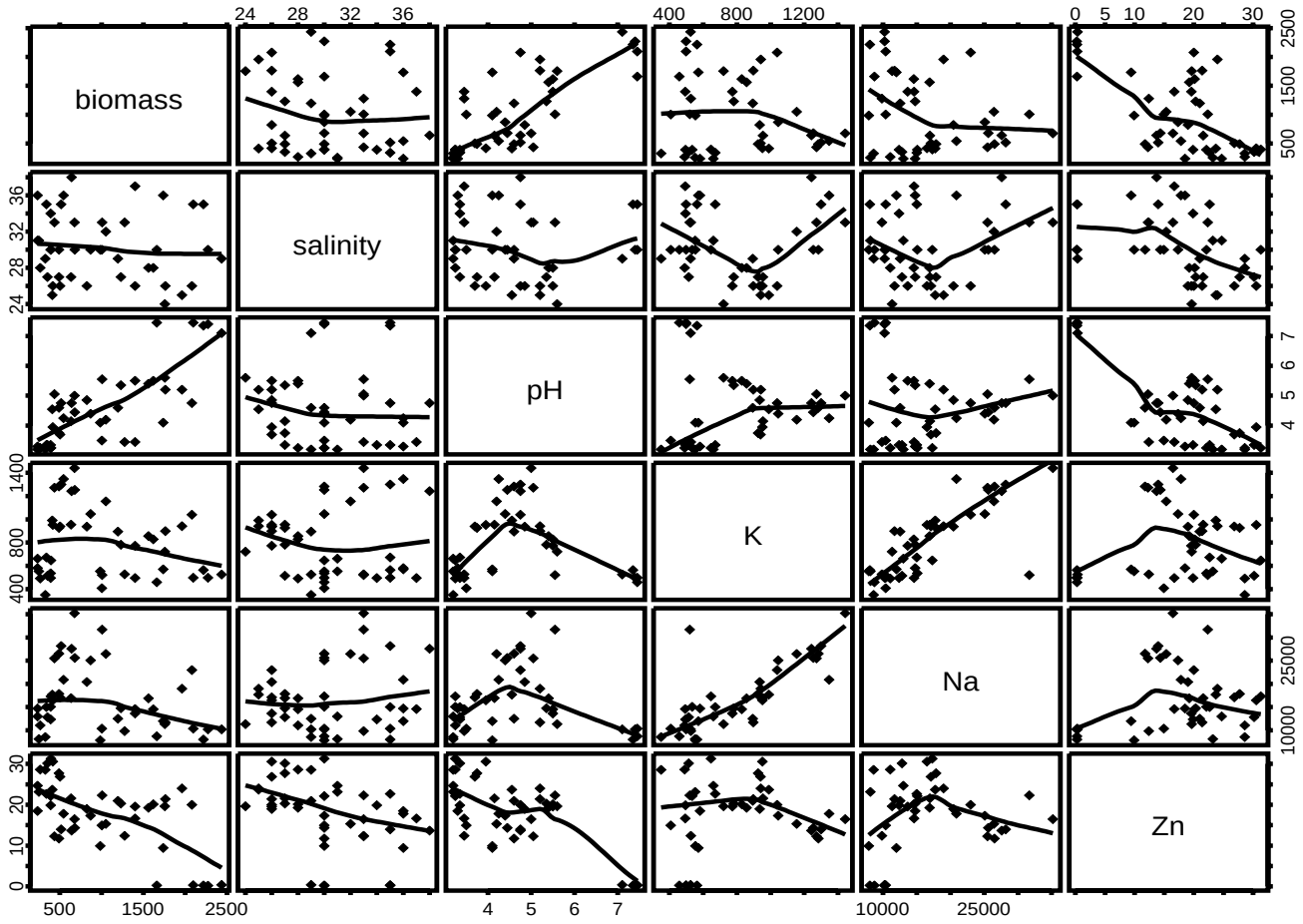
(b) The answer to the first question is yes if X has full column rank. The answer is no if X does not have full column rank. Let $\underline{b}^* = (X^TX)^-X^T\underline{Y} + \underline{a}$ for some generalized inverse $(X^TX)^-$ of X^TX. Then, $X^TX\underline{b}^* = X^TX(X^TX)^-X^T\underline{Y} + X^TX\underline{a} = X^T\underline{Y} + X^TX\underline{a}$. This is a solution to the normal equations if $X^TX\underline{a} = 0$. If X does not have full column rank, the columns of X are linearly dependent and there is at least one $\underline{a} \neq 0$ for which $X\underline{a} = 0$. Consequently, $X^TX\underline{a} = 0$ for that $\underline{a}$. If X has full column rank, then X^TX is a non-singular matrix and there is no $\underline{a} \neq \underline{0}$ for which $X^TX\underline{a} = 0$. (Some students failed to provide a correct derivation and wrongly argued that $X^TX\underline{a} = 0$ only if X^TX is a matrix of zeros. Note that the diagonal elements of X^TX are sums of the squared elements in the columns of X which could not be zeros unless X is a matrix of zeros.)

(c) It is easy to show that $b = (X^TX)^{-1}X^T\underline{Y}$ is a solution to the normal equations. Suppose there exits another solution to the normal equations, let's say $\underline{b}^*$. We can write $\underline{b}^* = \underline{b} + \underline{a} = (X^TX)^{-1}X^T\underline{Y} + \underline{a}$, for some $\underline{a} \neq 0$. Since, rank(X^TX)=rank(X) and X has full column rank, X^TX also has full column rank and the columns of X^TX are linearly independent. Therefore, it is impossible to have $X^TX\underline{a} = 0$, for some $\underline{a} \neq 0$. Consequently, $\underline{b} = (X^TX)^{-1}X^T\underline{Y}$ is the unique solution to the normal equations.

7.(a)

```
# Enter the data into a data frame.

>biomass <- read.table("c:/st511/biomass.txt",header=T)
>biomass

#  Compute correlations and round the results
#  to four significant digits

> round(cor(biomass[-(1:2)]),4)

#  Create a scatterplot matrix

> splom(~biomass[,3:8],aspect="fill")

#  Create a scatterplot matrix with smooth
#  curves.  Unix users  should first use motif( )
#  to open a graphics wundow

points.lines <- function(x, y)
  {
```

```
points(x, y)
lines(loess.smooth(x, y, 0.90))
 }

par(din=c(7,7), pch=18, mkh=.15, cex=1.2, lwd=3)
pairs(biomass[ ,-(1:2)], panel=points.lines)
```



There seems to be a positive linear relationship between Y and pH and a negative linear relationship between Y and Zn. The other three variables appear to be at more weakly correlated with the response Y, but Y seems to be lower for larger values of K and Na. With the exception of one potentially influencial case, it appears that K and Na have a strong positive correlation. There appears to be a negative association between levels of pH and Zn in the soil.

(b) The model matrix has full column rank.

```
# Construct the response vector and model matrix

> Y<-as.matrix(biomass[ ,3])
> X<-as.matrix(biomass[,4:8])
> X0 <- rep(1, length(Y))
> X<-cbind(X0,X)

# Check the rank of the model matrix

> qr(X)$rank
[1] 6
```

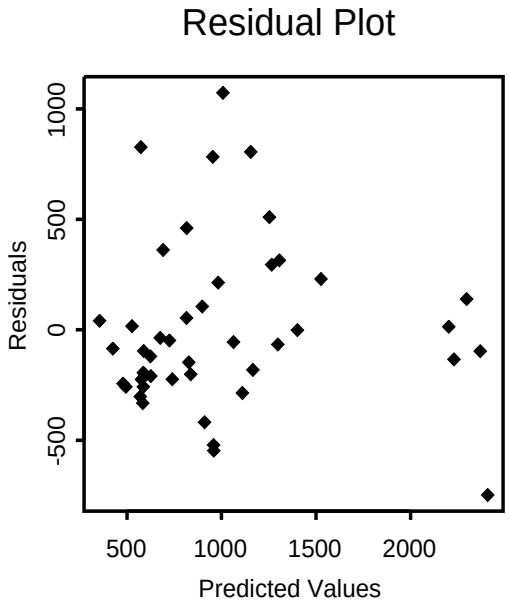
(c)
```
> b <- solve(t(X)%*%X) %*% t(X) %*% Y
>  b
                    [,1]
      X0  1.252455e+003
salinity -3.028491e+001
      pH  3.054880e+002
       K -2.852645e-001
      Na -8.665809e-003
      Zn -2.067677e+001
```

(d)
```
> yhat <- X %*% b
> yhat
        [,1]
 1  724.1509
 2  739.6501
 3  690.9358
 4  814.6608
 5 1063.5814
 6  957.8484
 7  527.0257
     .
     .
     .
42 1297.8306
43 1401.3330
44 1305.8814
45 1265.4053

> e <- Y - yhat
> e
           [,1]
 1   -48.150889
 2  -223.650127
 3   361.064216
 4    53.339248
 5   -55.581358
 6  -521.848358
 7    16.974252
         .
         .
         .
42   -65.830647
43    -1.332987
44   314.118581
45   294.594716
```

The motif( ) function shown below opens a graphics window on UNIX workstations. It should **not** be used in a Windows version of S-PLUS.

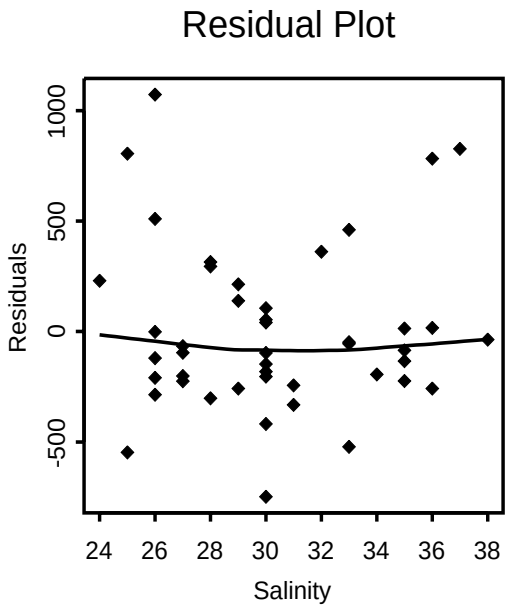
```
> motif()
# Plot residuals against fitted values #

# Specify plotting symbol and size of graph in inches.
# fin=c(w,h) specifies a plot that is w inches wide
#     and h inches high, not including labels
# pch=18  requests a filled diamond as a plotting
#          symbol
# mkh=b   requests plotting symbols that are  b
#          inches high
# cex=c   requests that the size of characters used to print
#         labels are c times the default for the printer
```

```
# mar     mar=c(5,5,4,2) defines the number of lines of
#          text allowed on each side of the figure,
#          starting on the bottom and moving clockwise.


> par(fin=c(8.0,8.0),pch=18,mkh=.1,cex=1.3,mar=c(5,5,4,2))

> plot(yhat,e,xlab="Predicted Values",
   ylab="Residuals", main="Residual Plot")
```

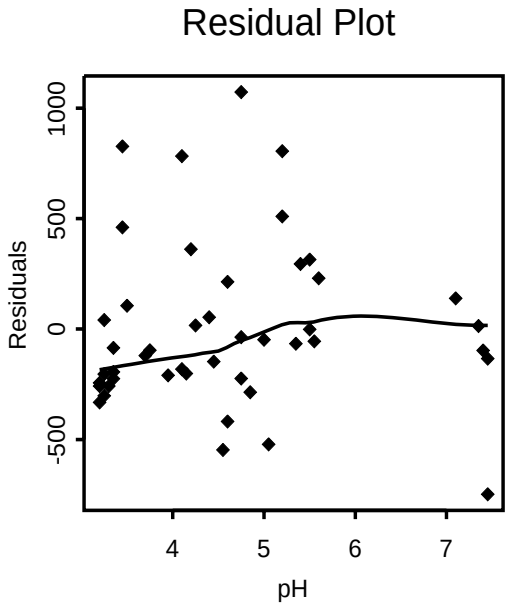


To some students this plot sugested that the assumption of constant variance does not hold. It is difficult to tell because there are only a few points on the right side of the plot. Another possibility is that the specified regression model is not a good approximation to the true model..

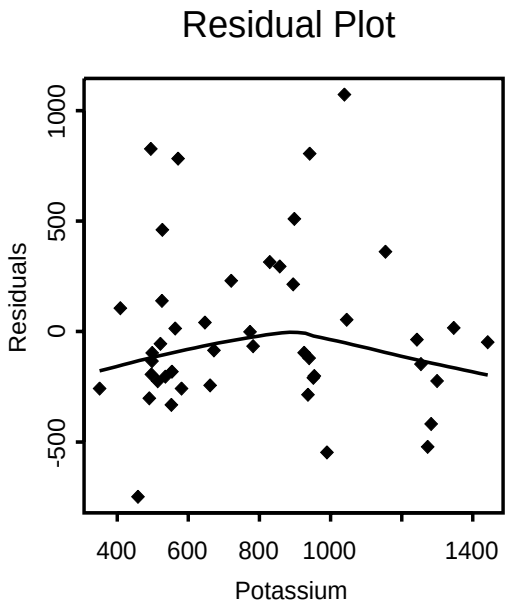
(e)
```
> plot(biomass$salinity,e,
+      xlab="Salinity",ylab="Residuals",main="Residual Plot")
> lines(loess.smooth(biomass$salinity, e, 0.90))
```



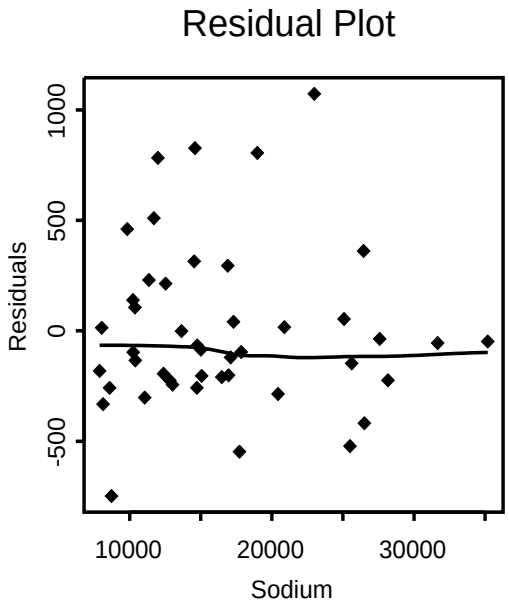

```
> plot(biomass$pH,e,
+      xlab="pH",ylab="Residuals",main="Residual Plot")
> lines(loess.smooth(biomass$pH, e, 0.90))
```



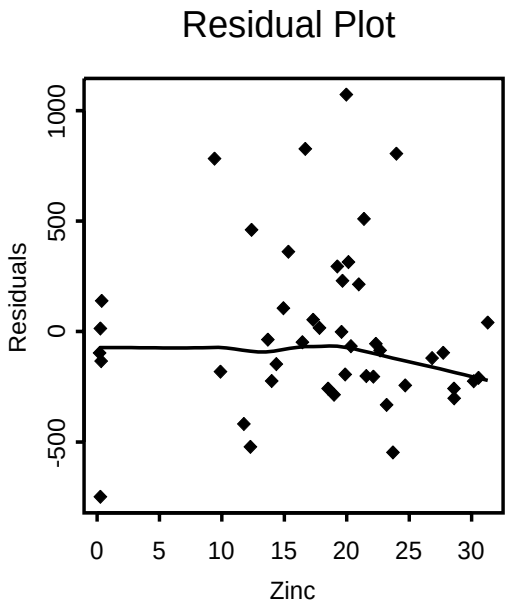

```
> plot(biomass$K,e,
+      xlab="Potassium",ylab="Residuals",main="Residual Plot")
> lines(loess.smooth(biomass$K, e, 0.90))
```




```
> plot(biomass$Na,e,
+      xlab="Sodium",ylab="Residuals",main="Residual Plot")
> lines(loess.smooth(biomass$Na, e, 0.90))
```

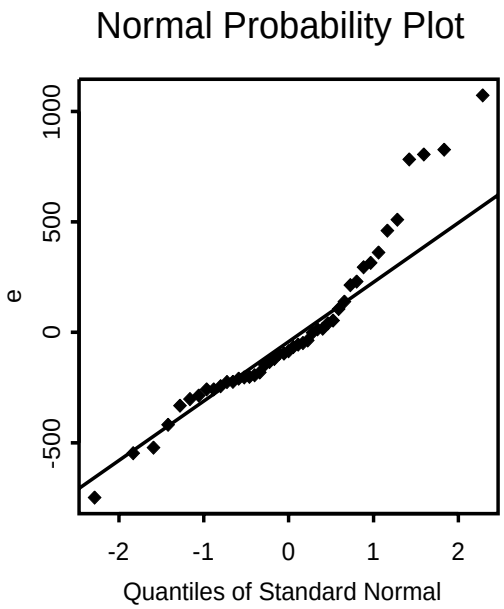



```
> plot(biomass$Zn,e,
+      xlab="Zinc",ylab="Residuals",main="Residual Plot")
> lines(loess.smooth(biomass$Zn, e, 0.90))
```



Based on the respective plots, the variance does not appear to be constant with respect to some of the explanatory variables, supporting our conclusion in part(d). There may be a curved effect of some of the variables on the mean response.

(f)
```
> qqnorm(e, main="Normal Probability Plot")
> qqline(e)
```



Based on this plot, the distribution of the random errors appears to be skewed to the right. This could be a result of non-homogeneous error variances or failure to incorporate curved relationships into the model.

(g)
```
> SSE <- crossprod(e,e)
> SSE
        [,1]
[1,] 6186015
> MSE <- SSE/(length(Y)-length(b))
> MSE
         [,1]
[1,] 158615.8
```

(h)
```
> Covb <- as.vector(MSE)*solve(t(X)%*%X)
> Covb
                     X0       salinity             pH
      X0  1524509.10834 -27237.9322767 -91209.1728455
salinity   -27237.93227    577.4385457   1276.4382277
      pH   -91209.17286   1276.4382280   7723.0703457
       K     -134.26839      2.5498036      3.7084102
      Na        7.40833     -0.1655202     -0.3417675
      Zn   -16373.54466    257.3248227   1104.7023266

                     K           Na             Zn
      X0 -134.268394262  7.408329862 -1.637354e+004
salinity    2.549803557 -0.165520186  2.573248e+002
      pH    3.708410204 -0.341767495  1.104702e+003
       K    0.121317254 -0.004495673  1.000216e+000
      Na   -0.004495673  0.000253753 -8.119987e-002
      Zn    1.000215850 -0.081199866  2.266258e+002

#  Compute standard errors

> stderrb <- sqrt(diag(Covb))
```

```
> coef <- c("Intercept","Salinity", "pH", "K", "Na", "Zn")
> heading <- c("Estimate","Std. Error")
> tempb <-  cbind(b, stderrb)
> dimnames(tempb) <- list(coef, heading)
> round(tempb,4)
            Estimate Std. Error
Intercept  1252.4546  1234.7101
 Salinity   -30.2849    24.0300
       pH   305.4880    87.8810
        K    -0.2853     0.3483
       Na    -0.0087     0.0159
       Zn   -20.6768    15.0541

# Write the estimates and standard errors to a file

> write.matrix(tempb,file="c:/courses/st511/hw/tempb.out")
```

(i) In Splus, a dataframe is a tow dimensional object that is similar to a matrix in appearance, both have rows and columns. All entries in a matrix must be of the same type, i.e., all entries are double precision numeric values, or all entries are character values. The columns of a data frame, on the other hand, correspond to set of vectors of the same length but possibly of *different* types, 1.e., one column can be a vector of character values and another column can be a vetor of numeric values.