Homework4

Problem 1

From the question, we choose L and K, and to maximize $0.05L^{2/3}K^{1/3}$ such that $L + K \leq 100000$.

```
labor = function(L){
  K = (100000-12*L)/15
  machines = 0.05*(L^(2/3))*(K^(1/3))
  return (-machines)
}
S=optim(10,labor,method="L-BFGS-B")
S$value
```

```
## [1] -204.6684
```

```
S$par
```

```
## [1] 5555.556
```

Thus, the maximum number of machines it can produce are 204. ## Problem 2 The file homework4stocks.csv contains historical monthly returns for 27 companies. The first row contains stock names, and the first column contains the dates. For each company, calculate the estimated mean return and the estimated variance of return. Then calculate the estimated correlations between the companies' returns.

Find a portfolio that achieves an expected monthly return of at least 1% and minimizes portfolio variance. What are the fractions invested in each stock? What are the portfolio's estimated mean, variance, and standard deviation? Assume no short selling is allowed.

```
library(quadprog)
m=c(mandv[2,])
s=c(mandv[1,])
rho = cor_stock
covMat=diag(s) %*% rho %*% diag(s)

Dmat=2*covMat
dvec=rep(0,27)
Amat = matrix(rep(0,810),27)
Amat[,1] = m
Amat[,2] = c(rep(1,27))
Amat[,3] =c(rep(-1,27))
Amat[,4:30] = diag(1,27,27)
bvec=c(0.01,1,-1,rep(0,27))


S=solve.QP(Dmat,dvec,Amat,bvec)

S$solution
```

```
##  [1] -1.022911e-17  9.061097e-19 -2.819441e-18 -7.240527e-18 -2.312358e-17
##  [6] -3.387596e-17 -1.492791e-16 -5.292693e-17 -8.201727e-18  8.552208e-18
## [11]  1.137924e-17 -5.401488e-18  0.000000e+00  3.369193e-17 -1.855199e-17
## [16]  1.310882e-01 -8.605708e-18  7.421209e-19  3.418563e-01 -8.070447e-18
## [21]  2.527138e-01 -9.199921e-18 -2.741344e-17  5.222670e-02  2.716464e-02
## [26] -1.064547e-17  1.949503e-01
```

```r
S$value
```

```
## [1] 2.364023e-06
```

What are the portfolio's estimated mean, variance, and standard deviation?

```r
crossprod(S$solution, m)
```

```
##      [,1]
## [1,] 0.01
```

```r
crossprod(S$solution, s)
```

```
##             [,1]
## [1,] 0.002395405
```

```r
sqrt(crossprod(S$solution, s))
```

```
##            [,1]
## [1,] 0.04894287
```

Problem 3

The file 'variable_selection.csv' contains observations of variables y, x1, x2, and x3. Here, y is the dependent variable. We want to choose a linear model that uses at most two independent variables such that the sum of squared residuals is minimized. This can be formulated as a constrained quadratic programming problem.

```r
variable = read.csv("variable_selection.csv")
l1 = lm(y~x1, data = variable)
l2 = lm(y~x2, data = variable)
l3 = lm(y~x3, data = variable)
l4 = lm(y~x1+x2, data = variable)
l5 = lm(y~x1+x3, data = variable)
l6 = lm(y~x2+x3, data = variable)
sum(l1$residuals^2)
```

```
## [1] 7901.299
```

```r
sum(l2$residuals^2)
```

```
## [1] 878.8358
```

```r
sum(l3$residuals^2)
```

```
## [1] 8575.636
```

```r
sum(l4$residuals^2)
```

```
## [1] 26.19087
```

```r
sum(l5$residuals^2)
```

```
## [1] 7860.089
```

```r
sum(l6$residuals^2)
```

```
## [1] 878.1811
```

L4 seems to be the best model.

Problem 4

1. Formulate a quadratic programming problem whose solution will yield the current flowing through each resistor.

Choose: x_{12},x_{13},x_{23},x_{24},x_{34}

minimize $x_{12}^2 + 4x_{13}^2 + 6x_{23}^2 + 12x_{24}^2 + 3x_{34}^2$

Subject to:

x_{12} + x13 = 710 x24 + x34 = 710 x12 - x23 - x24 = 0 x13 + x23 - x34 = 0

- Choose

x_{12},x_{13},x_{23},x_{24},x_{34}

- to minimize

$$x_{12}^2 + 4x_{13}^2 + 6x_{23}^2 + 12x_{24}^2 + 3x_{34}^2$$

- subject to

$$\begin{aligned} x_{12} + x_{13} &= 710 \\ x_{24} + x_{34} &= 710 \\ x_{12} - x_{23} - x_{24} &= 0 \\ x_{13} + x_{23} - x_{34} &= 0 \\ x_{12}, x_{13}, x_{23}, x_{24}, x_{34} &\geq 0 \end{aligned}$$

2. Use R to determine the current flowing through each resistor.

```
Dmat=2*matrix(c(1,0,0,0,0,0,4,0,0,0,0,0,6,0,0,0,0,0,12,0,0,0,0,0,3),5,5)
dvec=rep(0,5)
Amat = matrix(rep(0,45),5)
Amat[,1] = c(1,1,0,0,0)
Amat[,2] = c(0,0,0,1,1)
Amat[,3] = c(1,0,-1,-1,0)
Amat[,4] =c(0,1,1,0,-1)
Amat[,5:9]=diag(1,5,5)
bvec=c(710,710,rep(0,7))
S=solve.QP(Dmat,dvec,Amat,bvec,meq = 4)
S$solution
```

```
## [1] 371.3846 338.6154 163.8462 207.5385 502.4615
```

```
S$value
```

```
## [1] 2031911
```

Answer: 371.3846, 338.6154, 163.8462, 207.5385, 502.4615 will yield the current flowing shown above.

Problem 5

You will need to calculate the following:

- Actual Point Spread = Home Team Score – Visiting Team Score

- Predicted Spread = Home Team Rating – Visitor Team Rating + Home Team Advantage
- Prediction error = Actual Point Spread – Predicted Point Spread

You will also need to normalize the ratings. To do this, you set the actual average of the ratings to be 85 (this is somewhat arbitrary but based on the well-known Sagarin rating system).

What do these ratings mean? If two teams had ratings of 82 and 91, then the second team would be predicted to win by 9 points if the game was played on a neutral field.

```
nfl = read.csv(file="nflratings.csv", header=FALSE)
#Actual Point Spread
nfl$APS = nfl$V4-nfl$V5

rating = function(ratings){
  error = sum((nfl$APS - (ratings[nfl$V2]-ratings[nfl$V3]+ ratings[33]))^2)
  return (error)
}

nfl_ratings = c(rep(85,32),1)

S = optim(nfl_ratings, rating)

#optimal team ratings
S$par
```

```
##  [1] 85.430978 88.180000 91.245592 83.988501 87.159197 83.228365 82.631829
##  [8] 81.087756 91.762413 87.561464 76.775562 91.670425 84.099863 90.794124
## [15] 81.655778 78.384316 86.941679 91.799659 92.631536 91.905018 86.280875
## [22] 91.790590 78.905130 90.942557 87.223204 70.469942 91.480059 87.572041
## [29] 79.524789 80.738464 83.579456 82.958342  2.476405
```

The result is shown above, the home advantage is 2.476.