Homework 4

Brett Scroggins

Due 3/9/2018

Problem 1

1. $100000 = 15k + 12L$

$\hookrightarrow L = \dfrac{100000 - 15K}{12}$

$f(L) = .05\,L^{2/3}\,K^{1/3}$

The company will be able to create about 204 printers

```r
machines = function(K){

  L = (100000-15*K)/12
  machines = 0.05*L^(2/3)*K^(1/3)
  return(-machines)

}

optim(200, machines, method='BFGS')
```

```
## $par
## [1] 2216.799
##
## $value
## [1] -204.6681
##
## $counts
## function gradient
##       11       10
##
## $convergence
## [1] 0
##
## $message
## NULL
```

Problem 2

2.

$$A = \overset{A^T}{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \text{mean}(1) & \text{mean}(2) & \cdots & \text{mean}(28) \end{bmatrix}} \overset{X}{\begin{bmatrix} s1 \\ s2 \\ \vdots \\ s28 \end{bmatrix}} \geq \overset{b}{\begin{pmatrix} 1 \\ -1 \\ .01 \end{pmatrix}}$$

$d = (0, 0, \ldots, 0)$ $\quad$ $D = 2 \cdot (\text{covariance matrix})$

$\boxed{\text{Percent Return} = 1\%, \quad \text{Variance} = 0.0031, \quad \text{Std. Dev} = 0.0533}$

```
stocks = read.csv('/Users/brettscroggins/Downloads/homework4stocks.csv')

# Get info from the stocks data set
num = ncol(stocks)
mean = colMeans(stocks[,2:num])
std = sqrt(apply(stocks[,2:num],2,var))

corr = cor(stocks[,2:num], use = 'pairwise.complete.obs')
cov = cov(stocks[,2:28], use='pairwise.complete.obs')


# Create the A, b, d, and D components
A_mat = matrix(c(rep(1,27), diag(27), mean),27)
b_vec = c(1,rep(0,27),.01)
d_vec = rep(0,27)
```

```
D_mat = 2*cov

# Solution
sol2 = solve.QP(D_mat, d_vec, A_mat, b_vec)
perc_return = sum(sol2$solution*mean)
perc_return
```

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

```
variance = sum(sol2$solution*(std^2))
variance
```

```
## [1] 0.003101297
```

```
stnddev = sum(sol2$solution*std)
stnddev
```

```
## [1] 0.05330648
```

Problem 3

3.

Regression	Sum of Square Residuals
$y = \beta_1 x_1$	7901.299
$y_2 = \beta_2 x_2$	878.8358
$y_3 = \beta_3 x_3$	8575.636
$y_4 = \beta_1 x_1 + \beta_2 x_2$	26.19087
$y_5 = \beta_1 x_1 + \beta_3 x_3$	7860.089
$y_6 = \beta_2 x_2 + \beta_3 x_3$	878.1811

The regression that best fits the data (has the lowest sum of square residuals) is $y_4 = \beta_1 x_1 + \beta_2 x_2$

```r
vars = read.csv('/Users/brettscroggins/Downloads/variable_selection.csv')

lm1 = lm(vars$y ~ vars$x1)
lm2 = lm(vars$y ~ vars$x2)
lm3 = lm(vars$y ~ vars$x3)
lm4 = lm(vars$y ~ vars$x1 + vars$x2)
lm5 = lm(vars$y ~ vars$x1 + vars$x3)
lm6 = lm(vars$y ~ vars$x2 + vars$x3)

r1 = sum(resid(lm1)^2)
r1
```

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

```r
r2 = sum(resid(lm2)^2)
r2
```

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

```r
r3 = sum(resid(lm3)^2)
r3
```

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

```r
r4 = sum(resid(lm4)^2)
r4
```

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

```r
# r4 is lowest with 26.19 sum of square residuals
r5 = sum(resid(lm5)^2)
r5
```

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

```r
r6 = sum(resid(lm6)^2)
r6
```

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

Problem 4

4. • $I_1 + I_4 \le 710$

$I_1 \ge I_6 + I_{12} \rightarrow -I_1 + I_6 + I_{12} \le 0$

$I_3 \le I_4 + I_6 \rightarrow I_3 - I_4 - I_6 \le 0$

• Min: $1I_1^2 + 3I_3^2 + 4I_4^2 + 6I_6^2 + 12I_{12}^2$

$d = (0,0,0,0,0)$

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}$$

$$\underset{A^T}{\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 & 0 \end{bmatrix}} \underset{x}{\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \end{bmatrix}} \le \underset{b}{\begin{bmatrix} 710 \\ 0 \\ 0 \end{bmatrix}}$$

$$\boxed{I_1 = 371.4 \quad I_3 = 502.5 \quad I_4 = 338.6 \quad I_6 = 163.8 \quad I_{12} = 207.5}$$

```r
D_mat = matrix(0,5,5)
D_mat[1,1] = 1*2
D_mat[2,2] = 3*2
D_mat[3,3] = 4*2
D_mat[4,4] = 6*2
D_mat[5,5] = 12*2
d_vec = rep(0,5)
A_mat = matrix(c(1,0,1,0,0,-1,0,0,1,1,0,1,-1,-1,0),5,3)
b_vec = c(710,0,0)

solve.QP(D_mat,d_vec,A_mat,b_vec)
```

```
## $solution
## [1] 371.3846 502.4615 338.6154 163.8462 207.5385
##
## $value
## [1] 2031911
##
## $unconstrained.solution
## [1] 0 0 0 0 0
##
## $iterations
## [1] 4 0
##
## $Lagrangian
## [1] 5723.692 4980.923 3014.769
##
## $iact
## [1] 1 2 3
```

Problem 5

5. To normalize, the parameters from the solution were transformed the following way:

normalized score = parameter + (85 - mean (parameters))

- The ratings yield a score for the "home field advantage" effect of all 32 NFL teams.
- Two teams with ratings of 82 and 91 does NOT yield that them playing on a neutral field would result in a 9 point win for team 91.

What this reveals is that team 91 is above league average at home while team 82 is below average.

```r
nfl_ratings = read.csv('/Users/brettscroggins/Downloads/nflratings.csv', header = FALSE)

nfl_ratings$spread = nfl_ratings$V4 - nfl_ratings$V5

error = function(scores){

  num = nrow(nfl_ratings)
  pred_spread = 0

  for(i in 1:num){
    pred_spread = pred_spread + (nfl_ratings$spread[i] - (scores[nfl_ratings[i,2]] - score
  }
  return(pred_spread)
}
```

```r
start = rep(0,33)
sol5 = optim(start, error, method='BFGS')
normalized = sol5$par[1:32] + (85 - mean(sol5$par[1:32]))
normalized = c(normalized, sol5$par[33])
normalized
```

```
##  [1] 84.522358 89.841612 92.745183 83.089265 88.760069 79.812553 87.543497
##  [8] 76.886322 92.120642 85.636517 70.503351 92.255064 86.984740 90.862198
## [15] 78.439387 76.888502 86.615294 92.064543 96.123163 95.629714 85.099246
## [22] 93.148915 75.033602 90.958388 86.641868 67.720162 92.605720 85.241670
## [29] 74.731408 79.170973 82.187631 80.136444  2.172819
```