MATH 668 Homework 1 Solutions

1. The following R commands can be used to obtain the answers.

```r
require(HistData)
```

```
## Loading required package: HistData
```

```r
attach(GaltonFamilies)
mean(childHeight)
```

```
## [1] 66.74593
```

(a) The average height of all children is 66.74 inches.

```r
mean(gender=="female")
```

```
## [1] 0.4850107
```

(b) The percentage of female children is 48.50%.

```r
max(childHeight)
```

```
## [1] 79
```

(c) The height of the tallest child is 79 inches.

```r
sum(children==4)/4
```

```
## [1] 31
```

(d) The number of families with exactly 4 children is 31.

```r
tallest.mother.height=max(mother)
GaltonFamilies[mother==tallest.mother.height,]
```

```
##     family father mother midparentHeight children childNum gender
## 571    128   68.7   70.5           72.42        2        1   male
## 572    128   68.7   70.5           72.42        2        2 female
##     childHeight
## 571        71.0
## 572        61.7
```

(e) The heights of the children in the family with the tallest mother are 71 and 61.7 inches.

```r
mean((childHeight>father)&(childHeight>mother))
```

```
## [1] 0.238758
```

(f) The percentage of children that are taller than both of their parents is 23.88%.

```r
mean(childHeight[(gender=="female")&(mother>=68)])
```

```
## [1] 66.46667
```

(g) The average height of females whose mother was at least 68 inches tall is 66.47 inches.

```r
detach(GaltonFamilies)
```

2.(a) Since

$$\begin{aligned}\mathbf{H}\left(\alpha_1\mathbf{H}_1^- + \ldots + \alpha_m\mathbf{H}_m^-\right)\mathbf{H} &= \mathbf{H}(\alpha_1\mathbf{H}_1^-)\mathbf{H} + \ldots + \mathbf{H}(\alpha_m\mathbf{H}_m^-)\mathbf{H} \\ &= \alpha_1\mathbf{H}\mathbf{H}_1^-\mathbf{H} + \ldots + \alpha_m\mathbf{H}\mathbf{H}_m^-\mathbf{H} \\ &= \alpha_1\mathbf{H} + \ldots + \alpha_m\mathbf{H} \\ &= (\alpha_1 + \ldots + \alpha_m)\mathbf{H},\end{aligned}$$

$\alpha_1\mathbf{H}_1^- + \ldots + \alpha_m\mathbf{H}_m^-$ is a generalized inverse of $\mathbf{H}$ if and only if $\alpha_1 + \ldots + \alpha_m = 1$.

(b) Since $\mathbf{H}_1^- = \begin{pmatrix} 1/4 & 1/4 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\mathbf{H}_2^- = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 0 & 0 \\ -1 & 0 & -2 \end{pmatrix}$, and $\mathbf{H}_3^- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 1/3 & -1/3 \end{pmatrix}$ are generalized inverses of $\mathbf{H}$,

$$\alpha_1\mathbf{H}_1^- + \alpha_2\mathbf{H}_2^- + \alpha_3\mathbf{H}_3^- = \begin{pmatrix} \frac{1}{4}\alpha_1 + 2\alpha_2 & \frac{1}{4}\alpha_1 & 3\alpha_2 \\ -\frac{1}{2}\alpha_1 & \frac{1}{2}\alpha_1 + \frac{2}{3}\alpha_3 & +\frac{1}{3}\alpha_3 \\ -\alpha_2 & \frac{1}{3}\alpha_3 & -2\alpha_2 - \frac{1}{3}\alpha_3 \end{pmatrix}$$

is a generalized inverse of $\mathbf{H}$ if $\alpha_1 + \alpha_2 + \alpha_3 = 1$. So, one example with no zero elements is

$$4\mathbf{H}_1^- + 3\mathbf{H}_2^- - 6\mathbf{H}_3^- = \begin{pmatrix} 7 & 1 & 9 \\ -2 & -2 & -2 \\ -3 & -2 & -4 \end{pmatrix}.$$

3. First, we read the data into R.

```
O=rbind(
c(255, 54, 71, 87, 88, 93,255),
c( 54,255,255,255,255,255, 97),
c( 71,255,255,255,255,255,101),
c( 87,255,255,255,255,255,103),
c( 88,255,255,255,255,255,108),
c( 93,255,255,255,255,255,115),
c(255, 97,101,103,108,115,255))
```

The eigenvalues and eigenvectors can be obtained with the following commands.

```
E=eigen(O)
l=E$val;l
```

```
## [1]  1.373685e+03  4.151384e+02  1.289019e+01  1.256982e-13  2.620790e-14
## [6] -5.587959e-01 -1.615523e+01
```

```
v=E$vec;v
```

```
##            [,1]       [,2]        [,3]          [,4]          [,5]
## [1,] -0.2049720  0.6922375 -0.45765380  0.000000e+00  0.000000e+00
## [2,] -0.4186751 -0.1949149  0.53020254  3.521904e-01 -2.284790e-01
## [3,] -0.4219260 -0.1603606  0.09424771 -7.566207e-01  3.906608e-01
## [4,] -0.4246705 -0.1305774 -0.39000992  4.970539e-02 -4.997859e-01
## [5,] -0.4257125 -0.1211512 -0.21599687  5.223998e-01  6.629717e-01
## [6,] -0.4277085 -0.1019515 -0.10019312 -1.676749e-01 -3.253676e-01
## [7,] -0.2452864  0.6441876  0.54014312  1.249001e-14  6.938894e-16
##             [,6]       [,7]
## [1,] -0.02098782  0.5185617
## [2,]  0.09260833  0.5663812
## [3,]  0.20737922  0.1388647
## [4,]  0.54645723 -0.3156338
## [5,] -0.07064908 -0.2000308
```

```
## [6,] -0.79938550 -0.1537420
## [7,]  0.07316486 -0.4772307
```

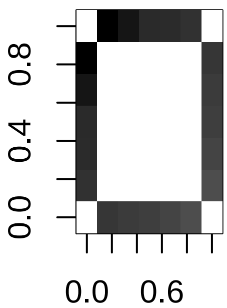
So, the two largest eigenvalues are 1373.685 and 415.1384. Then, we compute **A** as follows.

```
A=l[1]*v[,1]%*%t(v[,1])+l[2]*v[,2]%*%t(v[,2])
A
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
## [1,] 256.64467  61.87156  72.71689  82.04861  85.05086  91.13034 254.18759
## [2,]  61.87156 256.56357 255.63722 254.80571 254.64227 254.23665  88.94556
## [3,]  72.71689 255.63722 255.22106 254.82905 254.80547 254.68418  99.28166
## [4,]  82.04861 254.80571 254.82905 254.81563 254.91253 255.03615 108.17127
## [5,]  85.05086 254.64227 254.80547 254.91253 255.04781 255.24942 111.04319
## [6,]  91.13034 254.23665 254.68418 255.03615 255.24942 255.60953 116.85023
## [7,] 254.18759  88.94556  99.28166 108.17127 111.04319 116.85023 254.92156
```

The following figure compares the two images.

```
image(t(O[7:1,]),col=gray((0:255)/255))
```




```
image(t(A[7:1,]),col=gray((0:255)/255))
```