1. Over a 5-year period regular samples of fishermen on 28 lakes in Wisconsin were asked to report the time they spent fishing and how many of each type of game fish they caught. Their responses were then converted to a catch rate per hour for

x_1 = Bluegill $\qquad$ x_2 = Black crappie

x_3 = Smallmouth bass $\qquad$ x_4 = Largemouth bass

The estimated correlation matrix

$$\mathbf{R} = \begin{pmatrix} 1.0000 & .4919 & .2636 & .4653 \\ .4919 & 1.0000 & .3127 & .3506 \\ .2636 & .3127 & 1.0000 & .4108 \\ .4653 & .3506 & .4108 & 1.0000 \end{pmatrix}$$

is based on a sample of about 120 (there were a few missing values).

(a) Obtain the principal component solution for a factor model with $k = 1$.

(b) Obtain the maximum likelihood solution for a factor model with $k = 1$. Are the principal component and maximum likelihood solutions consistent with each other?

(c) Obtain the principal component solution for a factor model with $k = 2$, and rotate your solution. Interpret each factor.

2. The data file `AirPollution.csv` contains $n = 42$ measurements on air pollution variables recorded at 12:00 noon in the Los Angeles area on different days:

x_1 = Wind $\qquad$ x_2 = Solar radiation $\qquad$ $x_3 = NO_2$ $\qquad$ $x_4 = O_3$

Compute the sample correlation matrix.

(a) Obtain the principal component solution to a factor model with $k = 1$.

(b) Find the maximum likelihood estimates of $\mathbf{Q}$ and $\boldsymbol{\Psi}$ for $k = 1$.

(c) Compare the factorization obtained by the principal component and maximum likelihood methods.

(d) Perform a varimax rotation of the principal component solution to a factor model with $k = 2$. Interpret the results.