CS 635: Tools and Environments for Optimization
Homework 9: Quadratic Programming Exercises

Michael C. Ferris, Computer Sciences Department

March 26, 2009

Submit this assignment electronically using the instructions on the course web page. The assignment name is hw9 and you should hand in exactly 4 files with the following names:

hw9-1.gms, hw9-2.gms, results.gdx, hw9-3.gms

You will need to download from the course webpage the files nfl2002.gdx and the file laplace.gdx for some of the data for these problems.

1. Consider the undirected graph G on n nodes with edge weights W. Define a diagonal matrix D by $D_{i,i} = \sum_j W_{i,j}$ and the Laplacian matrix $Q = D - W$. Suppose that the variables f are fixed to values y on some subset l of the nodes. Solve the problem

$$\min_f f'Qf \text{ s.t. } f_l = y_l$$

   The data for the problem can be found in laplace.gdx. Experiment with a variety of "qcp" solvers, and hard code the one you wish to use by

```
option qcp=xxxx;
```

   in your gams file. This has application to text and speech recognition.

   Add constraints that force $f_i \geq 0.5$, $i \notin l$ and resolve.

2. We examined the results of the 2002 NFL season, and aim to use regression to determine a rating for each team. The file "nfl2002.gdx" contains the margins of victory by the home team in each game. Of course, a negative "margin of victory" indicates that the visiting team won the game.

   (a) We aim to choose the team ratings to predict the results of all these games as well as possible, according to some loss function. We also aim to find the "average" home field advantage. If team j (the visiting team) has rating R_j, and team i (the home team) has rating R_i, and the home field advantage is H, then our prediction of the margin of victory for the home team is is $R_i + H - R_j$

   Using a sum-of-squares objective, applied to the difference between our predicted outcomes and the actual outcomes from the file "nfl2002.gdx", determine the ratings for each team and the home field advantage. Apply the following constraint

to the ratings: They must sum to 0. In the same GAMS file, use an l_1 objective function in place of the sum-of-squares to obtain a different set of ratings. Comment on the differences in the ratings obtained by these two methods.

Ensure that you put the results into the 'SSQ' and 'L1' columns of a parameter

```
set modrun / 'SSQ', 'L1', 'SSQ-MOD', 'L1-MOD' /;
results(teams,teamnames,*);
```

(b) Now modify the question in part (a) as follows: Use a separate "home field advantage" variable H_i for each team. Specifically, if team j (the visiting team) has rating R_j, and team i (the home team) has rating R_i, and the home field advantage is H_i, then our prediction of the margin of victory for the home team is is $R_i + H_i - R_j$. Use sum-of-squares and l_1 objectives to find the best-fit ratings and home field advantages.

The results should be put under the "MOD" columns of results.

According to these ratings, who would you expect to win if the packers (GB) were to visit the seahawks (Seattle)? Save your answer as 1 or 0 in the scalar packersWin into the file results.gdx along with the parameter "results".

3. Consider a set of m points $x_1, \ldots, x_m$ in R^2. Given a circle with center z and radius r, let

$$d_i(u) = \|z - x_i\|^2 - r^2$$

be an approximation to the true distance from the data point x_i to the circle:

$$|\,\|z - x_i\| - r|$$

.

Let $u = (z_1, z_2, r)$: we wish to solve

$$\min_u \sum_{i=1}^{m} d_i(u)^2$$

Formulate this problem as a quadratic program and solve for the points (1,1), (-3,2), (1,-5), (-2,4) in R^2.

Hint: Consider the change of variables, $y = (2z_1, 2z_2, r^2 - z^T z)$ and $b_i = (x_{i1}, x_{i2}, 1)$ so that $b_i^T y = 2z^T x_i - z^T z + r^2$, and the problem becomes

$$\min_y \sum_{i=1}^{m} (b_i^T y - x_i^T x_i)^2$$