# Math 77C Project 2 (Due May 31)

**Instructions**: This project will require some MATLAB code. It may therefore be convenient to answer the questions below and provide the accompanying MATLAB code in the same `.m` file. When turning in the project, please name your file `project2_yourlastname.m` and email it to `eesser@uci.edu`. If you are submitting multiple files, please zip them together in a file named `project2_yourlastname.zip`.

**1**. Suppose we are given Kobe Bryant's average points per playoff game for each of the previous ten NBA seasons and that these averages are normally distributed with variance $\sigma^2 = 10$ about an unknown mean $w$.

| **Year** | **Average points per game** |
|---|---|
| 2002 | 29.4 |
| 2003 | 26.6 |
| 2004 | 32.1 |
| 2005 | 24.5 |
| 2006 | 27.9 |
| 2007 | 32.8 |
| 2008 | 30.1 |
| 2009 | 30.2 |
| 2010 | 29.2 |
| 2011 | 22.8 |

Based on this, we want to predict Kobe's average points per game for the 2012 playoffs.

**1a**. Let $t$ be the vector of the observed points per game averages. The likelihood function for $w$ is

$$p(t|w) = \prod_{i=1}^{10} \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta(w-t_i)^2}{2}},$$

where $\beta = \frac{1}{\sigma^2}$. As a function of $w$, this product of Gaussians is again a Gaussian. What is the mean and variance?

**1b**. What $w$ maximizes the likelihood function?

**1c**. Now suppose we add a prior $p(w) = \sqrt{\frac{\alpha}{2\pi}} e^{-\frac{\alpha(w-c)^2}{2}}$ for $c = 25$ and $\alpha = .05$ based on the assumption that his average points per game for the playoffs is normally distributed with mean 25 and variance 20. In the same MATLAB figure, plot the likelihood $p(t|w)$, the prior $p(w)$ and the posterior distribution $p(t|w)p(w)$ all normalized to integrate to one.

**1d**. Find the maximum a posteriori (MAP) estimate of $w$ by taking minus the natural log of the posterior and then minimizing the resulting function with respect to $w$. How does this compare to the maximum likelihood (ML) estimate from 1b?

**1e**. Suppose the variance for our prior goes down and $\alpha$ increases to 1. What is the new MAP estimate of $w$? Plot the likelihood along with the new prior and posterior in the same MATLAB figure.

**2**.Suppose we are trying to fit a cubic polynomial to five points actually generated from the cubic polynomial $h(x) = 3 + 10x - 10x^2 + x^3$ except that one point is a severe outlier. Let the data points be given by

| $x$ | $y$ |
|---|---|
| 1 | $-1000$ |
| 2 | $-9$ |
| 6 | $-81$ |
| 9 | 12 |
| 10 | 103 |

**2a**. What are the coefficients $c$ of the best fit cubic in the least squares sense? Write this as a minimization problem of the form min $\frac{\beta}{2}\|Ac - y\|^2$.

**2b**. Now add a regularization term of the form $\frac{\alpha}{2}\|c\|^2$. What equation can you solve to minimize

$$\frac{\beta}{2}\|Ac - y\|^2 + \frac{\alpha}{2}\|c\|^2 \ ?$$

**2c**. Tabulate the best fit coefficients $c$ for $\beta = 1$ and $\alpha = 0, 1, 10, 100, 1000$

**2d**. In the same MATLAB figure, plot the data points, the function $h$, and all five best fit cubics under the different regularization assumptions.