Advanced Statistical Inference
Bayesian Linear Regression

1 Aims

- Implement the maximum likelihood estimate.
- Use this to explore uncertainty in parameter values.
- Make predictions.

2 LOO CV

Make sure you've completed the previous assignments.

If you want to look at the effect of increasing the polynomial order above 4, you'll need to rescale the x (Olympic year) values (do this before you calculate all the x^k stuff). The following code will do this (remember that you need to do the same thing to any test years too...)

```
x = male100(:,1);
t = male100(:,2);
mival = min(x); % First Olympic year
x = x - mival; % Subtract the First year
x = x./4; % Olympics are every four years

testx = [2012];
testx = testx - mival;
testx = testx/4;
```

3 Maximum likelihood

1. Using the ML estimators given in the lectures, compute $\widehat{\mathbf{w}}$ and $\widehat{\sigma^2}$ on the Olympic data for polynomials of first, second and third order. ($t = w_0 + w_1 x$, $t = w_0 + w_1 x + w_2 x^2$ and $t = w_0 + w_1 x + w_2 x^2 + w_3 x^3$).
2. Plot polynomial order versus $\widehat{\sigma^2}$ – what do you expect to see? What do you see?
3. Using your CV code from the previous exercise, perform a LOO CV analysis for maximum likelihood for polynomials of order 0 to 8.

4 Uncertainty in estimates

1. For the first order polynomial, compute $\mathsf{cov}\{\widehat{\mathbf{w}}\}$ (use $\widehat{\sigma^2}$ in place of σ^2):

$$\mathsf{cov}\{\widehat{\mathbf{w}}\} = \sigma^2 \left(\mathbf{X}^\mathsf{T}\mathbf{X}\right)^{-1}$$

2. Download `gausssamp.m` from the ASI collaborative space.
3. Sample 10 values of $\mathbf{w}$ from $\mathcal{N}(\widehat{\mathbf{w}}, \mathsf{cov}\{\widehat{\mathbf{w}}\})$ and plot them all the models with the data. Hints:

```
coW = ss*inv(X'*X); %ss is the estimate of sigma^2
samps = gausssamp(wHat,coW,10); % Arguments are mean, covariance and number of samples
samps = samps'; % Comes out the wrong way around
% Add code to plot the data
% And plotting the ith line
plot(x,X*samps(:,i),'r');
```

5 Uncertainty in predictions

1. Download and load `synthetic_data.mat` from the ASI collaborative space. It includes the variables:

```
x
t
testx
```

2. Pick a polynomial order between 1 and 7 (repeat with others if you have time)
3. Find $\widehat{\mathbf{w}}$ and $\widehat{\sigma^2}$
4. Compute the predictive means. If you create $\mathbf{X}_{\mathsf{new}}$ in exactly the same way as you created $\mathbf{X}$, you'll be able to do this for all test points in one operation:

$$t_{\mathsf{new}} = \mathbf{X}_{\mathsf{new}}\widehat{\mathbf{w}}$$

5. For each $\mathbf{x}_{\mathsf{new}}$ (each value of `testx` expanded into a vector, each row of $\mathbf{X}_{\mathsf{new}}$ etc etc), compute the predictive variance (you could do this in one operation but it will be easier to write a for loop):

$$\sigma^2_{\mathsf{new}} = \sigma^2 \mathbf{x}_{\mathsf{new}}^\mathsf{T}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{x}_{\mathsf{new}}$$

6. Use the following command to plot the data, predictive means and variances:

```
% Predictive means are called tnew, variances vnew.
% Both are vectors of the same size (and orientation!)
plot(x,t,'b.','markersize',25);
hold on
errorbar(testx,tnew,vnew)
```

7. What do you notice – is it what you expect?
8. Repeat with the data given in the file `synthetic_data_2.mat`. What do you notice?