10-601
Machine Learning

Problem Set 2
MATLAB tips

• Conventional wisdom: MATLAB *hates* loops
  – May be less of an issue with most recent versions
  – Ideally use matrix operations whenever possible

• Examples:

\[
A = \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\]
MATLAB tips

• Conventional wisdom: MATLAB *hates* loops
  – May be less of an issue with most recent versions
  – Ideally use matrix operations whenever possible

• Examples:

```
>> exp(A)

ans =

    2.7183    7.3891   20.0855
54.5982  148.4132  403.4288
```
MATLAB tips

• Conventional wisdom: MATLAB *hates* loops
  – May be less of an issue with most recent versions
  – Ideally use matrix operations whenever possible

• Examples:

```matlab
>> 10 + A
ans =
    11     12     13
    14     15     16

>> 2 ./ A
ans =
    2.0000       1.0000       0.6667
    0.5000       0.4000       0.3333
```
MATLAB tips

• Conventional wisdom: MATLAB *hates* loops
  – May be less of an issue with most recent versions
  – Ideally use matrix operations whenever possible

• Examples:

\[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 1
\end{bmatrix}
\]

\[
\begin{bmatrix}
3 \\
6
\end{bmatrix}
\]

\[
\text{ans} =
\]

\[
\text{ans} =
\]

\[
\text{mod}(A, 3) == 0
\]

\[
A(\text{mod}(A, 3) == 0)
\]
MATLAB tips

• Conventional wisdom: MATLAB *hates* loops
  – May be less of an issue with most recent versions
  – Ideally use matrix operations whenever possible

• Examples:

```
>> A(sum(A,2)>10,2:3)

ans =

    5    6
```
Logistic regression implementation

• In my code that solves for $w$, the only loop is the one that iterates until the weight vector has converged
• My code converges in $\sim40k$ iterations (35.6s) when
  – $\lambda = 0$
  – $w$ and $v$ initialized to $1$
  – $k = 0.5$

• Can use ‘tic’ and ‘toc’ to time your code
Logistic regression implementation

- Tracking the likelihood and norm of the difference in weight vectors at every iteration can be informative.
- L1 regularization, \( \lambda = 5 \)
Logistic regression implementation

- Tracking the likelihood and norm of the difference in weight vectors at every iteration can be informative
- L1 regularization, lambda = 5

```matlab
>> plot(norms(44300:44420))
```
Logistic regression implementation

• To test your code, observe what happens to the weight vector as you increase lambda

$$\text{hist}(\mathbf{w})$$
Bayesian linear regression

\[ y = w_0 + w_1 x \]

- **No observations**
- **1 observation**
- **2 observations**
- **20 observations**

Image from Bishop
We also discussed...

• Why the Gaussian prior on w in Bayesian linear regression is a conjugate prior, as well as how to compute the posterior
• How to compute information gain in a decision tree with continuous attributes