Scaling RL

Supervised Learning

Regression

1 handout: slides

asst 4 milestones were due

blog entries were due

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

Scaling Reinforcement Learning

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

Model known (solving MDP):

- value iteration
- policy iteration: compute U^{π} using
  - linear algebra
  - simplified value iteration
  - a few updates (modified PI)

Model unknown (RL):

- ADP using
  - value iteration
  - a few updates (eg, prioritized sweeping)
- Q-learning

Function Approximation

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

$$\hat{U}(s) = \theta_0 f_0(s) + \theta_1 f_1(s) + \theta_2 f_2(s) + \ldots$$

Function Approximation

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

$$\hat{U}(s) = \theta_0 f_0(s) + \theta_1 f_1(s) + \theta_2 f_2(s) + \ldots$$
$$\hat{U}(x, y) = \theta_0 + \theta_1 x + \theta_2 y$$

Function Approximation

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

$$\hat{U}(s) = \theta_0 f_0(s) + \theta_1 f_1(s) + \theta_2 f_2(s) + \ldots$$
$$\hat{U}(x,y) = \theta_0 + \theta_1 x + \theta_2 y$$

given sample u at $s = x, y$, want update to decrease error:

$$E = \frac{(\hat{U}(s) - u)^2}{2}$$

Function Approximation

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

$$
\begin{aligned}
\hat{U}(s) &= \theta_0 f_0(s) + \theta_1 f_1(s) + \theta_2 f_2(s) + \ldots \\
\hat{U}(x,y) &= \theta_0 + \theta_1 x + \theta_2 y
\end{aligned}
$$

given sample u at $s = x, y$, want update to decrease error:

$$
\begin{aligned}
E &= \frac{(\hat{U}(s) - u)^2}{2} \\
\theta_i &\leftarrow \theta_i - \alpha \frac{\delta E}{\delta \theta_i}
\end{aligned}
$$

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

$$
\begin{aligned}
\hat{U}(s) &= \theta_0 f_0(s) + \theta_1 f_1(s) + \theta_2 f_2(s) + \ldots \\
\hat{U}(x,y) &= \theta_0 + \theta_1 x + \theta_2 y
\end{aligned}
$$

given sample u at $s = x, y$, want update to decrease error:

$$
\begin{aligned}
E &= \frac{(\hat{U}(s) - u)^2}{2} \\
\theta_i &\leftarrow \theta_i - \alpha \frac{\delta E}{\delta \theta_i} \\
\theta_i &\leftarrow \theta_i - \alpha(\hat{U}(s) - u) \frac{\delta \hat{U}(s)}{\delta \theta_i}
\end{aligned}
$$

Function Approximation

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

$$\hat{U}(s) = \theta_0 f_0(s) + \theta_1 f_1(s) + \theta_2 f_2(s) + \ldots$$
$$\hat{U}(x, y) = \theta_0 + \theta_1 x + \theta_2 y$$

given sample u at $s = x, y$, want update to decrease error:

$$E = \frac{(\hat{U}(s) - u)^2}{2}$$
$$\theta_i \leftarrow \theta_i - \alpha \frac{\delta E}{\delta \theta_i}$$
$$\theta_i \leftarrow \theta_i - \alpha(\hat{U}(s) - u)\frac{\delta \hat{U}(s)}{\delta \theta_i}$$

in other words, the updates are:

$$\theta_0 \leftarrow \theta_0 - \alpha(\hat{U}(s) - u)$$
$$\theta_1 \leftarrow \theta_1 - \alpha(\hat{U}(s) - u)x$$
$$\theta_2 \leftarrow \theta_2 - \alpha(\hat{U}(s) - u)y$$

Break

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

- projects: proposals, talks, paper
- asst 4

Class Outline

Scaling RL
- Summary
- Approx U
- Break
- Class Outline

Supervised Learning

Regression

knowledge: specialized, propositional, FOL, STRIPS
reasoning: constraints, FOL, induction, planning

1. search: heuristics, CSPs, games
2. knowledge representation: FOL, resolution
3. planning: STRIPS, MDPs
4. learning: RL, supervised, unsupervised
5. KR with uncertainty: HMMs, Bayes nets

Scaling RL

Supervised Learning
- The Setting
- k-NN

Regression

Supervised Learning

The Setting

Scaling RL
Supervised Learning
- The Setting
- k-NN

Regression

labeled examples
hypothesis space
free parameters = degrees of freedom
classification vs regression
noise, overfitting

Using the k-Nearest Neighbors

Scaling RL
Supervised Learning
- The Setting
- k-NN
Regression

majority. $k = 1$ gives Voroni cells

$$d(a,b) = \sqrt{\sum_i (a_i - b_i)^2}$$

normalize dimensions (divide by $\sqrt{\frac{1}{N}\sum_i (x_i - \bar{x})^2}$)
weight by distance?

$+$: robust to noise, choose k by easy cross-validation
$-$: memory, kd-tree, irrelevant features, sparse data in high d

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

Regression

Regression

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$\hat{y} = \theta_0 f_0(x) + \theta_1 f_1(x) + \theta_2 f_2(x) + \ldots$$

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$\hat{y} = \theta_0 f_0(x) + \theta_1 f_1(x) + \theta_2 f_2(x) + \ldots$$

$$\hat{y} = \theta_0 + \theta_1 x$$

$$\hat{y} = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$

$$\hat{y} = \theta_0 + \theta_1 \sin x$$

On-line Regression

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$y = \theta f(x)$$

given sample x, y, want update to decrease $E = \frac{(\hat{y}-y)^2}{2}$:

On-line Regression

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$y = \theta f(x)$$

given sample x, y, want update to decrease $E = \frac{(\hat{y}-y)^2}{2}$:

$$\theta_i \quad \leftarrow \quad \theta_i - \alpha \frac{\delta E}{\delta \theta_i}$$

On-line Regression

Scaling RL

Supervised Learning

Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$y = \theta f(x)$$

given sample x, y, want update to decrease $E = \frac{(\hat{y}-y)^2}{2}$:

$$\begin{aligned}
\theta_i &\leftarrow \theta_i - \alpha \frac{\delta E}{\delta \theta_i} \\
\frac{\delta E}{\delta \theta_i} &= \frac{\delta}{\delta \theta_i} \frac{(\hat{y}-y)^2}{2} \\
&= (\hat{y}-y) \frac{\delta}{\delta \theta_i} (\hat{y}-y) \\
&= (\hat{y}-y) \frac{\delta}{\delta \theta_i} (\theta f(x) - y) \\
&= (\hat{y}-y) f(x)
\end{aligned}$$

On-line Regression

$$y = \theta f(x)$$

given sample x, y, want update to decrease $E = \frac{(\hat{y}-y)^2}{2}$:

$$\begin{aligned}
\theta_i &\leftarrow \theta_i - \alpha \frac{\delta E}{\delta \theta_i} \\
\frac{\delta E}{\delta \theta_i} &= \frac{\delta}{\delta \theta_i} \frac{(\hat{y}-y)^2}{2} \\
&= (\hat{y}-y) \frac{\delta}{\delta \theta_i} (\hat{y}-y) \\
&= (\hat{y}-y) \frac{\delta}{\delta \theta_i} (\theta f(x) - y) \\
&= (\hat{y}-y) f(x) \\
\theta &\leftarrow \theta - \alpha(\hat{y}-y) f(x)
\end{aligned}$$

The LMS Procedure

Scaling RL

Supervised Learning

Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$\begin{aligned} y &= \theta x \\ \theta &\leftarrow \theta - \alpha(\hat{y} - y)x \end{aligned}$$

The LMS Procedure

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

$$
\begin{aligned}
y &= \theta x \\
\theta &\leftarrow \theta - \alpha(\hat{y} - y)x
\end{aligned}
$$

for $x = \langle 0, x_1, x_2 \rangle$, the updates are:

$$
\begin{aligned}
\theta_0 &\leftarrow \theta_0 - \alpha(\hat{y} - y) \\
\theta_1 &\leftarrow \theta_1 - \alpha(\hat{y} - y)x_1 \\
\theta_2 &\leftarrow \theta_2 - \alpha(\hat{y} - y)x_2
\end{aligned}
$$

$\alpha \approx 1/N$? or $100/(100 + N)$? or 0.1?
Δ rule, LMS weight update, Adaline rule, Widrow-Hoff rule, Perceptron rule
converges if data are linear
perceptron in finite time if linearly separable

EOLQs

Scaling RL
Supervised Learning
Regression
- Regression
- On-line Regression
- LMS
- EOLQs

- What question didn't you get to ask today?
- What's still confusing?
- What would you like to hear more about?

Please write down your most pressing question about AI and put it in the box on your way out.
Thanks!