

An aerial photograph of a city, likely Madison, Wisconsin, taken from a high vantage point looking down at the city and the surrounding water. The sun is setting behind a hill in the background, casting a warm, golden glow over the entire scene. The city is densely packed with buildings, and the water is filled with numerous sailboats. A large, semi-transparent rectangular box is overlaid on the center of the image, containing the course title in large, bold, black text.

CS/ECE/ME 532 Matrix Methods in Machine Learning

Welcome!





The next three weeks...

- This week: finish Unit 5: hinge loss, SVM, stochastic GD
- Next week: start Unit 6: kernel methods
- The week after: finish Unit 6: neural networks + A4 review.
- Tuesday, May 7: Assessment 4

Due dates:

- Tomorrow: Assignment 8
- *Note*: two assignments are due in the last week of classes:
 - Wednesday, May 1: Assignment 9
 - Friday, May 3: Assignment 10

Activity 19



Loss functions

Classifying new data:

features \searrow weights \swarrow

$$\hat{y} = \text{sign}(\mathbf{x}^T \mathbf{w})$$

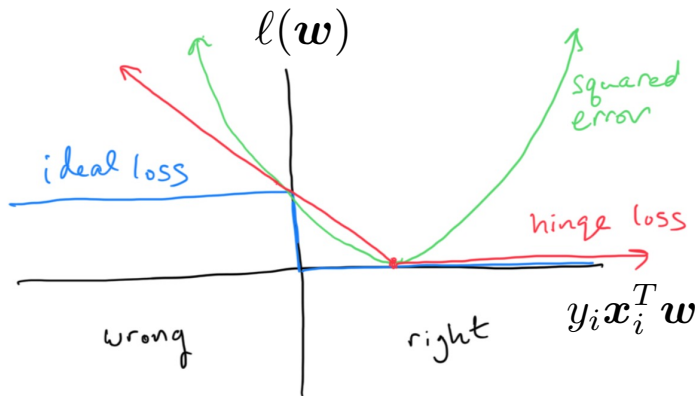
Training a classifier:

loss function \searrow

$$\min_{\mathbf{w}} \ell(\mathbf{w}) + \lambda r(\mathbf{w})$$

$\ell(\mathbf{w})$

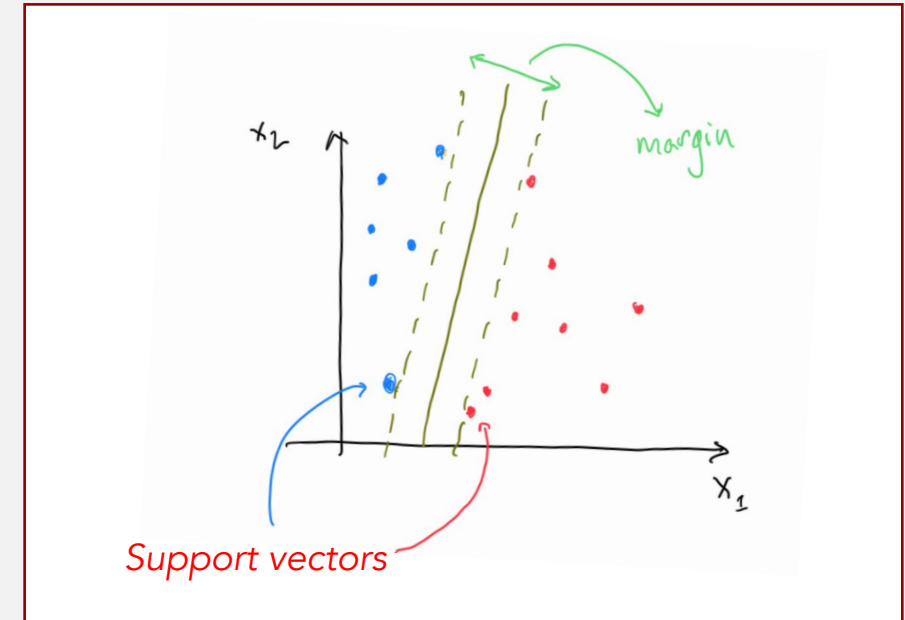
- squared error $\|\mathbf{X}\mathbf{w} - \mathbf{y}\|_2^2$
- ideal (0-1) loss $\sum_i \frac{1}{2} |y_i - \text{sign}(\mathbf{x}_i^T \mathbf{w})|$
- hinge loss $\sum_i (1 - y_i \mathbf{x}_i^T \mathbf{w})_+$
- logistic loss $\log(1 + e^{-y_i \mathbf{x}_i^T \mathbf{w}})$



$$y_i \mathbf{x}_i^T \mathbf{w}$$

positive when correct
negative when wrong

Support Vector Machines



maximize margin
s.t. correct classification

minimize $\|\tilde{\mathbf{w}}\|^2$
s.t. $y_i \mathbf{x}_i^T \mathbf{w} \geq 1$ for $i = 1, \dots$

For non-separable data: $\min_{\mathbf{w}} \sum_i (1 - y_i \mathbf{x}_i^T \mathbf{w})_+ + \lambda \|\tilde{\mathbf{w}}\|_2^2$