

Supervised Learning - Regularization

Adapting to overfitting

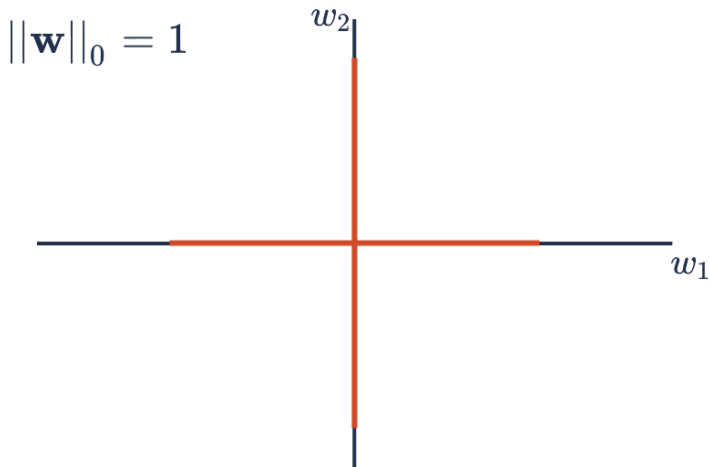
$\ell_p, \ell_{p,q}$ – norms

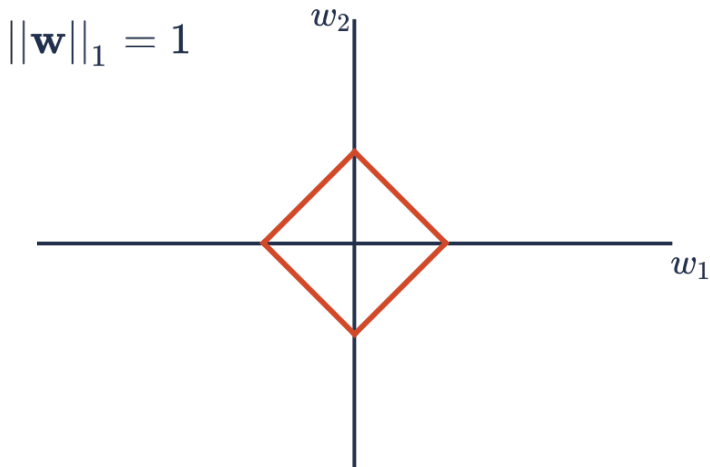
$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}}.$$

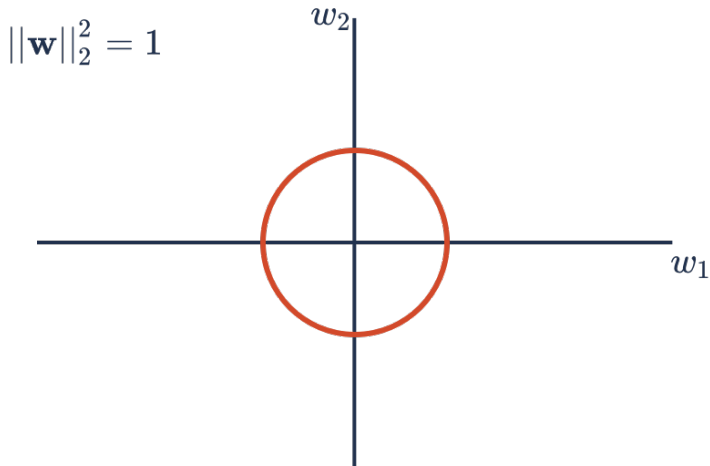
$$\|\mathbf{x}\|_{p,q} = \left(\sum_{j=1}^n \|\mathbf{x}_j\|_p^q \right)^{\frac{1}{q}}.$$

$$\|\mathbf{x}\|_{p,q} = \left(\sum_{j=1}^n \|\mathbf{x}_j\|_p^q \right)^{\frac{1}{q}} = \left(\sum_{j=1}^n \left(\sum_{i=1}^m x_{ij}^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}}.$$

$$||\mathbf{X}||_F = ||\mathbf{X}||_{2,2}.$$







ℓ_p -norms



$$p = \infty$$



$$p = 2$$



$$p = 1$$



$$0 < p < 1$$



$$p = 0$$

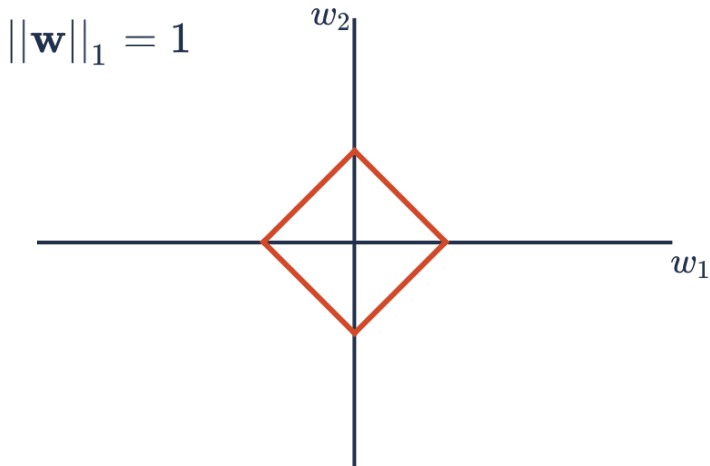
Linear Regression & Regularizations

Linear Regression

$$\min_{\mathbf{w}} \|\mathbf{y}^T - \mathbf{w}^T \mathbf{X}\|_2^2.$$

$$\min_{\mathbf{w}} \|\mathbf{y}^T - \mathbf{w}^T \mathbf{X}\|_2^2 + \alpha \|\mathbf{w}\|_1,$$

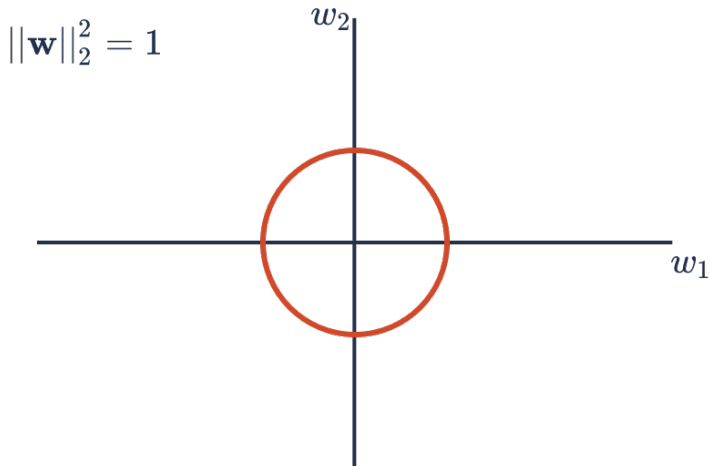
where α is a constant hyperparameter of our model.



Ridge Regression

$$\min_{\mathbf{w}} \|\mathbf{y}^T - \mathbf{w}^T \mathbf{X}\|_2^2 + \alpha \|\mathbf{w}\|_2^2,$$

where α is a constant hyperparameter of our model.



$$\min_{\mathbf{w}} \|\mathbf{y}^T - \mathbf{w}^T \mathbf{X}\|_2^2 + \lambda_1 \|\mathbf{w}\|_1 + \lambda_2 \|\mathbf{w}\|_2^2,$$

where λ_1, λ_2 are the coefficients for the ℓ_1 and ℓ_2 -norms respectively.

$$\min_{\mathbf{w}} \|\mathbf{y}^T - \mathbf{w}^T \mathbf{X}\|_2^2 + \alpha \rho \|\mathbf{w}\|_1 + \frac{\alpha(1 - \rho)}{2} \|\mathbf{w}\|_2^2.$$

where α is the normalization coefficient and ρ is a balancing ratio between the ℓ_1 and ℓ_2 -norms.

Questions

These slides are designed for educational purposes, specifically the CSCI-470 Introduction to Machine Learning course at the Colorado School of Mines as part of the Department of Computer Science.

Some content in these slides are obtained from external sources and may be copyright sensitive. Copyright and all rights therein are retained by the respective authors or by other copyright holders. Distributing or reposting the whole or part of these slides not for academic use is HIGHLY prohibited, unless explicit permission from all copyright holders is granted.