

# MACHINE LEARNING I — EXERCISE 4

Sacha Sokoloski, Sebastian Damrich

03 Nov 2025. Deadline: 10 Nov 2025 6pm

**Task 1** Let  $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$  be the singular value decomposition of the predictor matrix in linear regression, and  $\mathbf{y}$  be the response vector. In lecture 2, we computed that

$$\hat{\mathbf{y}} = \mathbf{U}\mathbf{U}^\top \mathbf{y}.$$

Now consider ridge regression with regularization parameter  $\lambda$ . Derive the analogous formula for  $\hat{\mathbf{y}}_\lambda$  and for  $\hat{\boldsymbol{\beta}}_\lambda$ . Explain how ridge regression affects the different elements of  $\hat{\boldsymbol{\beta}}_\lambda$  compared to the MSE solution  $\hat{\boldsymbol{\beta}} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^\top \mathbf{y}$ . You may assume that  $\mathbf{X}$  has full rank  $p + 1$ .

*Hint:* The derivation of  $\hat{\mathbf{y}} = \mathbf{U}\mathbf{U}^\top \mathbf{y}$  in lecture 2 might be helpful. Start with  $\hat{\boldsymbol{\beta}}_\lambda$ .

**Task 2** Consider linear regression with regularization penalty  $\lambda\|\boldsymbol{\beta}\|^4$ . Write down the gradient of the regularized loss function, and the gradient update step.

**Task 3** Let the prior for a scalar random variable be  $\mathcal{N}(\mu_1, \sigma_1^2)$  and the likelihood  $\mathcal{N}(\mu_2, \sigma_2^2)$ . Derive the mean  $\mu$  and the variance  $\sigma^2$  of the posterior distribution.

*Hint:* The posterior will again be a Gaussian distribution! You do not need to consider the normalization terms. Rewrite the product of two exponentials of squared difference as exponential of a single squared difference.

**Task 4** Consider simple (non-nested) 10-fold cross-validation on a dataset with  $n = 1000$ .

- What is the size of each training set?
- What is the size of each test set?
- Over the entire course of the cross-validation procedure, how many samples out of 1000 will serve as test samples?

*Hint:* Just provide the answers for the first two questions and a brief explanation for the third.