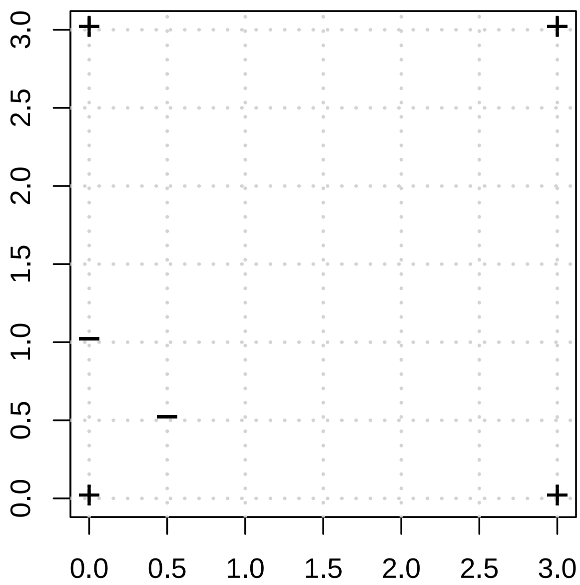
Exercise 1: SVM – Support Vectors and Separating Hyperplane

The primal optimization problem for the two-class soft margin SVM classification is given by

$$\min_{\theta,\theta_0,\zeta^{(i)}} \quad \frac{1}{2}||\theta||^2 + C\sum_{i=1}^{n}\zeta^{(i)}$$

$$\text{s.t. :} \quad y^{(i)}(\theta^\top \mathbf{x}^{(i)} + \theta_0) \geq 1 - \zeta^{(i)},$$

$$\zeta^{(i)} \geq 0, \quad \forall i = 1, \ldots, n.$$



(a) Add the decision boundary to the figure for $\hat{\theta} = (1,1)^T, \hat{\theta}_0 = -2$. (NB: This is the approximate optimum for $C = 10$)

(b) Identify the coordinates of the support vector(s) and compute the values of their slack variables $\zeta^{(i)}$.

(c) Compute the Euclidean distance of the non-margin-violating support vector(s) (i.e. support vectors that are located on the margin hyperplanes) to the decision boundary.

(d) What needs to be changed in the plot such that a hard margin SVM results into the same decision boundary?

Exercise 2: SVM – Optimization

Write your own stochastic subgradient descent routine to solve the soft-margin SVM in the primal formulation.

Hints:

- Use the regularized-empirical-risk-minimization formulation, i.e., an optimization criterion without constraints.
- No kernels, just a linear SVM.
- Compare your implementation with an existing implementation (e.g., `kernlab` in R). Are your results similar? Note that you might have to switch off the automatic data scaling in the already existing implementation.