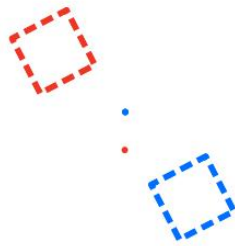
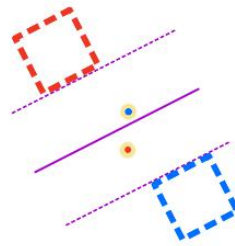


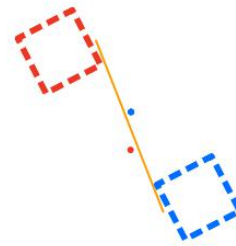
(c) Intuitively, when γ increases, the soft-margin SVM will behave much more like a common SVM due to the high cost of allowing $\xi_i > 0$. For example, for linear separable datasets, a very large γ will lead to an SVM that exactly separate two classes (hence the empirical risk is zero), while a small γ may allow some points to be in the buffer zone or even to lie on the wrong side of the affine hyperplane (hence the empirical risk is greater than zero). Therefore, it seems that the empirical risk will decrease for increasing γ . The following plots show an example of this circumstance.



Dashed lines represents multiple data points.

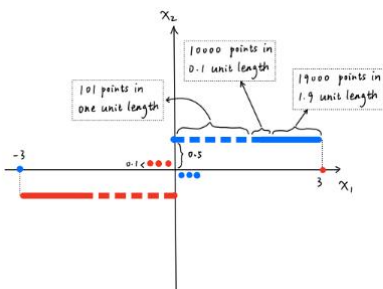


Soft-margin SVM with low λ (empirical risk = $\frac{2}{n}$)

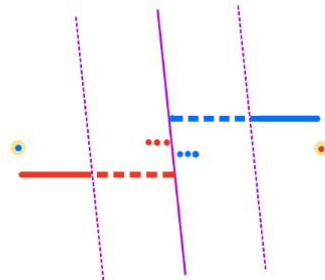


Soft-margin SVM with high λ (empirical risk = 0)

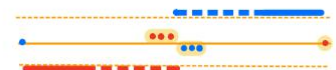
However, there are actually chances when some data points lie too far away from the main part of the data. In this case, the soft-margin SVM with high γ may be highly governed by these outliers, otherwise the ξ_i 's corresponding to these outliers may be too large. For example, if we have two outliers, the hyperplane (in \mathbb{R}^2) of SVM with high γ may turned out to be close to the line determined by these two points. This may lead to misclassification of multiple points lying just around this line. But a soft-margin SVM with very small γ can probably allow the outliers to be misclassified or even lie far beyond the margin at the wrong side, because even if the ξ_i 's may be quite large in this case, the tiny γ guarantees a relatively small total loss. Hence the empirical risk may increase for increasing γ . The following plots show an example of this circumstance.



Dashed lines represent multiple data points. Solid lines represent multiple data points with higher density.



Soft-margin SVM with $\lambda = 0.0001$. Empirical risk = $\frac{2}{n}$.



Soft-margin SVM with $\lambda = 1$. Empirical risk = $\frac{7}{n}$ (or maybe $\frac{6}{n}$, depending on the specific training procedure).

Therefore, we cannot conclude whether the empirical risk will increase or decrease for increasing γ .

PS: I verified the second case with a dataset generated by R and two soft-margin SVMs in Python. See https://github.com/trTime-ww/Soft_Margin_SVM_Question if you want to verify it yourself.