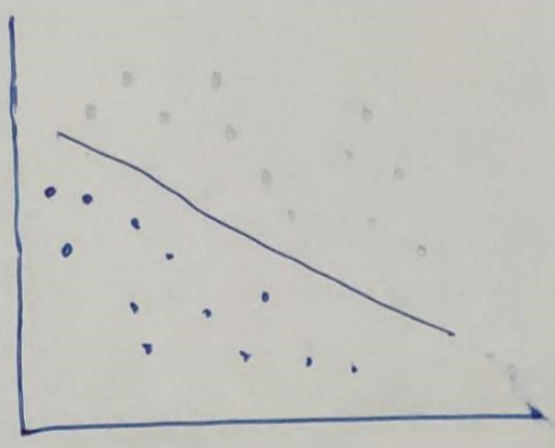
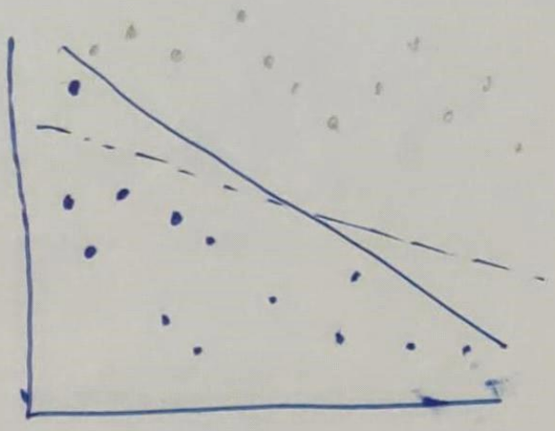
Support vector classifiers

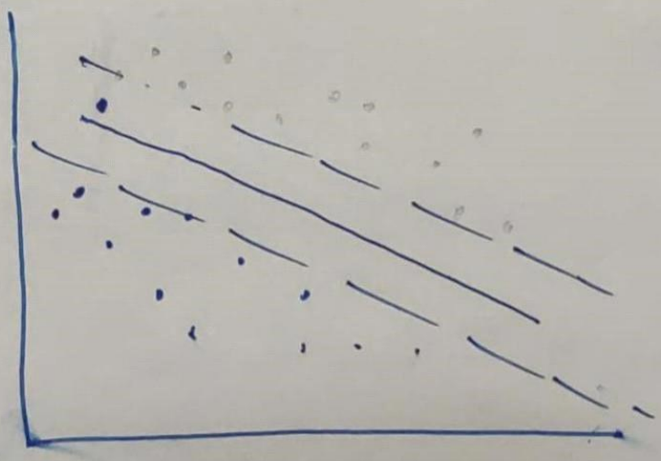
Suppose the classes are not separable directly, then





The deviation of the maximum margin classifier will be a lot; which decreases the robustness.

So we need to choose soft-margin classifier. where we have a certain allowance of errors that occurs.



Now the support vectors are given a range of tolerance.

So, now the equation changes

$$\underset{\beta_0, \beta_1, \beta_2, \cdots \beta_n \\ \epsilon_1, \epsilon_2, \epsilon_3 \cdots \epsilon_n}{\text{maximize}} \; M$$

$$\text{Subject to} \sum_{j=1}^{p} \beta_j^2 = 1,$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}, \cdots \beta_p x_{ip}) \geq M(1 - \epsilon_i),$$

$$\epsilon_i \geq 0 \sum_{i=1}^{n} \epsilon_i \leq C.$$

C is the non-negative tuning parameters. M is the width of the margin, but now slack variables are added. $(\epsilon_1, \epsilon_2, \epsilon_3 \cdots)$.