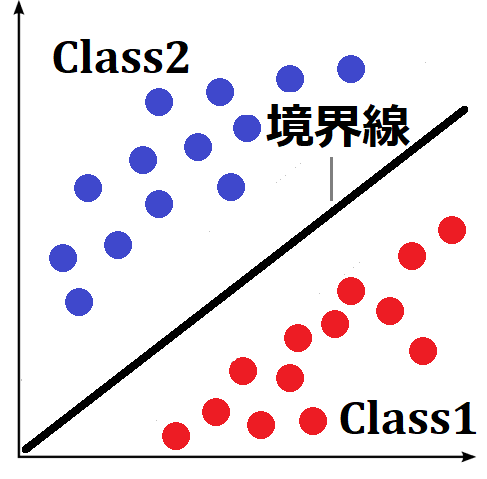
# 【SVM　サポートベクターマシーン】

高い汎化能力をもつ線形識別器の１つ。

SVMはクラスを分ける境界線を引くことを課題として行う。

そのための仕組みとして「マージン最大化」というアイデアを使う。

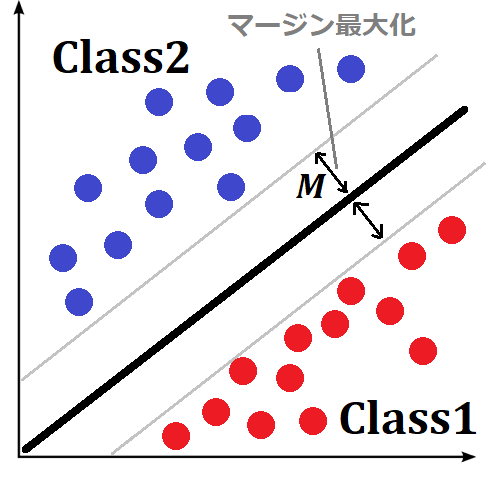


## 【マージン最大化】

マージンとは、「境界線」と「2つのクラス」の間の距離Mのこと。

マージン最大化は、距離Mが最大となるように境界線を決定すること。

距離が大きくなるほど、より多くの未学習データの判別が可能になる。



## 【ハードマージンSVMとソフトマージンSVM】

マージン最大化によって、分類するSVMだが現実の問題にはノイズが混じっているため、綺麗に分類出来ないことが多い。そのため、誤分類が発生してしまう。

その誤分類を許容するのがソフトマージンSVM。許容しないのがハードマージンSVM。

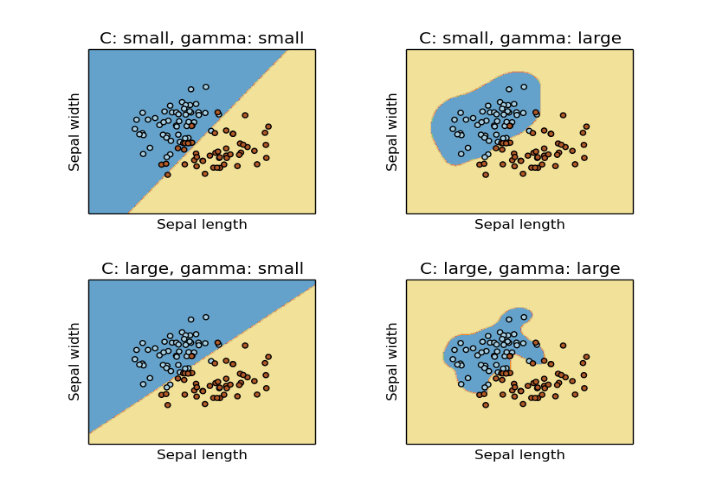
## 【sckit-learnでの実装】

**svm**.SVC(gamma, C)

|  |  |
| --- | --- |
| パラメータ | 説明 |
| gamma | ガウシアンカーネル（rbf, poly, sigmoid等）  のパラメータ。値が多きいほど複雑な境界になる。 |
| C | どれだけ誤分類を許容するかのパラメータ。  値が小さいほど誤分類を許容する。  （ソフトマージンSVMとなる） |
| kernel | カーネル関数の種類を選ぶ。デフォルトは  rbgとなる。  linear(線形カーネル:これを使うと単純な  線形になるが処理速度は速い。)  poly(多項式カーネル)  sigmoid（シグモイド関数を使ったモノ）  などを指定できる。 |

以下はrbfを使ったSVMでCとgammaの大きさと分類についてのデモになる。

※C は 2^-5 , 2^15に設定。　gammaは2^-15, 2^3に設定。



よって、SVMを使用する時は、Cとgammaのハイパーパラメータの設定がキーになる。