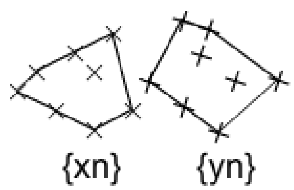
PRML 演習 4.1

@daimatz

(演習 4.1) データ $\{\mathbf{x}_n\}$ の集合が与えられ，凸包 (convex hull) とは以下の式で与えられるすべての点の集合であると定義することができる．

$$\mathbf{x} = \sum_n \alpha_n \mathbf{x}_n$$

ここで $\alpha \geq 0$ であり，$\sum_n \alpha_n = 1$ である．第 2 のデータ $\{\mathbf{y}_n\}$ の集合とそれに対応する凸包を考える．もしすべてのデータ $\mathbf{x}_n$ に対し $\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{x}_n + w_0 > 0$ を満たし，すべてのデータ $\mathbf{y}_n$ に対し $\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{y}_n + w_0 < 0$ を満たすベクトル $\hat{\mathbf{w}}$ とスカラー w_0 が存在するなら，定義によりこれら 2 つのデータの集合は線形分離可能である．それらの凸包が重なる場合，2 つのデータの集合は線形分離可能ではないことを示せ．また，逆に 2 つのデータの集合が線形分離可能な場合，それらの凸包が重ならないことを示せ．



図を書けば直感的には明らかだが，きちんと証明する．

$\{\mathbf{x}_n\}$ と $\{\mathbf{y}_m\}$ の凸包が重なるならば，

$$\mathbf{z} = \sum_n \alpha_n \mathbf{x}_n = \sum_m \beta_m \mathbf{y}_m$$

$$(\alpha_n > 0,\ \sum_n \alpha_n = 1,\quad \beta_m > 0,\ \sum_m \beta_m = 1)$$

を満たすような点 $\mathbf{z}$ が存在する．このとき $\{\mathbf{x}_n\}$ と $\{\mathbf{y}_m\}$ が線形分離可能であると仮定すると，

$$\begin{gathered}\text{すべてのデータ } \mathbf{x}_n \text{ に対し } \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{x}_n + w_0 > 0 \\ \text{かつ} \\ \text{すべてのデータ } \mathbf{y}_m \text{ に対し } \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{y}_m + w_0 < 0\end{gathered}$$

を満たす $\hat{\mathbf{w}}$ と w_0 が存在する．よってそのような $\hat{\mathbf{w}}$ と w_0 に対しては

$$\begin{aligned}\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{z} + w_0 &= \sum_n \alpha_n \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{x}_n + w_0 \\ &\quad (\sum_n \alpha_n = 1 \text{ より}) \\ &= \sum_n \alpha_n (\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{x}_n + w_0) \\ &\quad (\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{x}_n + w_0 > 0,\ \alpha_n \geq 0,\ \exists n[\alpha_n > 0] \text{ より}) \\ &> 0\end{aligned}$$

かつ

$$
\begin{aligned}
\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{z}+w_0 &= \sum_m \beta_m \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{y}_m + w_0 \\
&\quad (\sum_m \beta_m = 1 \text{ より}) \\
&= \sum_m \beta_m (\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{y}_m + w_0) \\
&\quad (\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{y}_m + w_0 < 0,\ \beta_m \geq 0,\ \exists m[\beta_m > 0] \text{ より}) \\
&< 0
\end{aligned}
$$

となるのでこの 2 つの式が等しいことに矛盾する．したがって，$\{\mathbf{x}_n\}$ と $\{\mathbf{y}_m\}$ の凸包が重なるならば $\{\mathbf{x}_n\}$ と $\{\mathbf{y}_m\}$ は線形分離可能でない．

　また $\{\mathbf{x}_n\}$ と $\{\mathbf{y}_n\}$ が線形分離可能であると仮定し，$\mathbf{z}$ がそれらの凸包の重なりの部分に存在することを考えても，同じ矛盾が起こるのでそのような $\mathbf{z}$ は存在しないことがわかる．