Least squares classification

1) Binary classification

$\mathcal{D} = \{(x_i, y_i)_{i=1}^{N}\}$ $\quad x_i^T \in \mathbb{R}^p \;;\; y_i \in \{+1, -1\}$ for $i = 1, 2, \ldots, N$

$y = f(x)$ where $f : \mathbb{R}^p \longrightarrow \{+1, -1\}$: unknown classifier/ functional relationship.

__Objective:__ To "guess" f !! (denote this guess/model/ estimated classifier $\hat{f}$)

Let $f_1, f_2, \ldots, f_k$ be basis fns from $\mathbb{R}^p \to \mathbb{R}$

$$\tilde{f} = \alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_k f_k$$

$$\hat{\alpha} = \underset{\alpha \in \mathbb{R}^k}{\operatorname{argmin}} \sum_{i=1}^{N} (y_i - \tilde{f}(x_i))^2 \leftarrow$$

$$\hat{f}(x_i) = \operatorname{sign}(\tilde{f}(x_i)) \quad \forall\, i = 1, 2, \ldots, N$$

$\operatorname{sign}(a) = +1$ if $a > 0$
$\qquad\quad\;\; = -1$ if $a < 0$

		Predicted case	
		+1	-1
True cases	+1 (T)	True positive	false negative
	-1 (F)	False positive	True negative Negatives

1) $\hat{f}$ is the "estimated classifier".

2) $\hat{f}$ is coming from $\tilde{f}$.

Can we interprete actual $\tilde{f}$ values as confidence in predicting the class??

3) Generalize to many classes.