LINEAR CLASSIFIERS

Important subclass of classification models.

Definition: If discriminant(s) $f_k(\mathbf{x})$ can be written as affine linear function(s) (possibly through a rank-preserving, monotone transformation g):

$$g(f_k(\mathbf{x})) = \boldsymbol{w}_k^\top \mathbf{x} + b_k,$$

we will call the classifier **linear**.

- $\boldsymbol{w}_k$ and b_k do not necessarily refer to parameters $\boldsymbol{\theta}_k$, although they often coincide; discriminant simply must be writable in an affine-linear way
- reasons for the transformation is that we only care about the position of the decision boundary

LINEAR DECISION BOUNDARIES

We can also easily show that the decision boundary between classes i and j is a hyperplane. For every $\mathbf{x}$ where there is a tie in scores:

$$
\begin{aligned}
f_i(\mathbf{x}) &= f_j(\mathbf{x}) \\
g(f_i(\mathbf{x})) &= g(f_j(\mathbf{x})) \\
\boldsymbol{w}_i^\top \mathbf{x} + b_i &= \boldsymbol{w}_j^\top \mathbf{x} + b_j \\
(\boldsymbol{w}_i - \boldsymbol{w}_j)^\top \mathbf{x} + (b_i - b_j) &= 0
\end{aligned}
$$

This represents a **hyperplane** separating two classes: