86-631/42-631 Neural Data Analysis
Lecture 12: Estimation and Classification III - LDA

1) Linear Discriminant Analysis: a special case of MAP estimation
    a. The boundary equation
    b. LDA in 1D
    c. LDA in 2D
2) **Slides: Optical Imaging of Neuronal Populations During Decision-Making (Briggman et al.)**

Linear Discriminant Analysis

Suppose you are trying to predict which of several stimuli produced a given set of data. Further suppose that the data (conditioned on the stimuli) follow a multivariate Gaussian distribution:

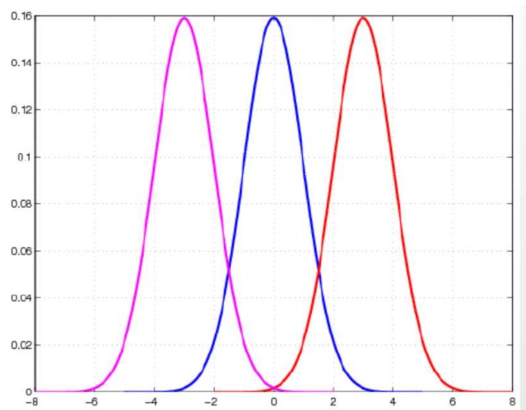
$$f(\vec{x}|C) = \frac{1}{(2\pi)^{k/2} det(\mathbf{\Sigma_C})^{1/2}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_C)^T \mathbf{\Sigma}_C^{-1}(\vec{x}-\vec{\mu}_C)}$$

Further suppose that for all of these stimuli, the covariance matrices will be the same for both classes, that is:

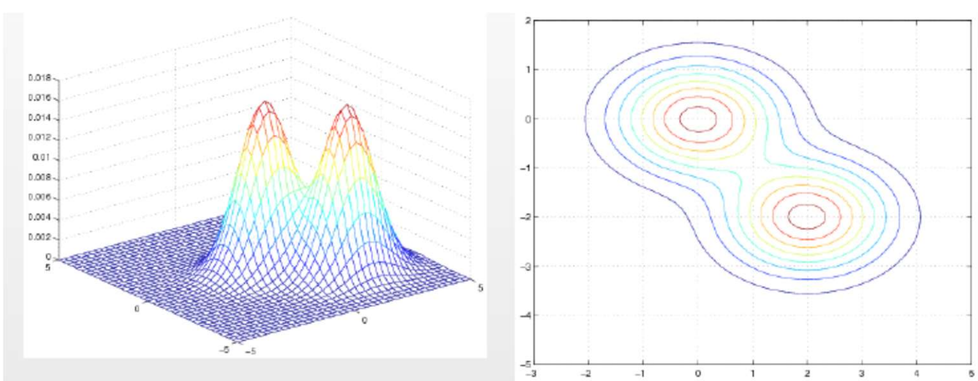
$$\Sigma_C = \Sigma \ \forall \ C.$$

In other words, the conditional distributions of the data for each category are just shifted versions of one another:

In 1 D



In 2D



The MAP estimate of C for this case would be:

$$\hat{C} = \arg\max_{C}\{\pi_C f(\vec{x}|C)\}$$

$$\hat{C} = \arg\max_{C}\left\{\frac{\pi_C}{(2\pi)^{k/2}det(\mathbf{\Sigma})^{1/2}}e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_C)^T\mathbf{\Sigma^{-1}}(\vec{x}-\vec{\mu}_C)}\right\}$$

(Here, π_C is the prior probability of class C. This is commonly used notation in the literature.)

Now, if Σ doesn't depend on C, the denominator of this function is a constant for all C, and can be ignored – it doesn't affect the solution!

Similarly, we've mentioned that solving for the argmax of a function is the same as solving for the argmax of the log of the function. So, for this problem, the solution for C reduces to solving:

$$\hat{C} = \arg\max_{C}\left\{\log\pi_C - \frac{1}{2}(\vec{x}-\vec{\mu}_C)^T\mathbf{\Sigma^{-1}}(\vec{x}-\vec{\mu}_C)\right\}$$

Expanding the second term yields:

$$\hat{C} = \arg\max_{C}\left\{\log\pi_C - \frac{1}{2}\left(\vec{x}^T\mathbf{\Sigma^{-1}}\vec{x} - \vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{x} - \vec{x}^T\mathbf{\Sigma^{-1}}\vec{\mu}_C + \vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{\mu}_C\right)\right\}$$

Now, the first term in parentheses does not depend on C. So, again, it will not affect the solution. Thus, we can rewrite the equation as:

$$\hat{C} = \arg\max_{C}\left\{\log\pi_C - \frac{1}{2}\left(\vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{\mu}_C - \vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{x} - \vec{x}^T\mathbf{\Sigma^{-1}}\vec{\mu}_C\right)\right\}$$

Further, we can take advantage of a simple linear algebra trick, and realize that the last two terms are exactly the same:

$$\vec{x}^T\mathbf{\Sigma^{-1}}\vec{\mu}_C = \vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{x}$$

(They are transposes of one another, and they are both scalars!)

So, the solution for C becomes:

$$\hat{C} = \arg\max_{C}\left\{\vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{x} - \frac{1}{2}\vec{\mu}_C^T\mathbf{\Sigma^{-1}}\vec{\mu}_C + \log\pi_C\right\}$$

Note the term in brackets is linear in X! We will define the ***linear discriminant function*** $\delta(x)$ to be the inside of the brackets:

$$\delta_C(x) = \mu_C^T \mathbf{\Sigma^{-1}} x - \frac{1}{2}\mu_C^T \mathbf{\Sigma^{-1}} \mu_C + \log \pi_C$$

(where for convenience I've dropped the vector notation: it's understood that x is a vector.)

So, our solution for C is written as

$$\hat{C} = \arg\max_C \{\delta_C(x)\}$$

The decision boundary between two stimuli k and l are all of the data x that lead to $\delta_k(x) = \delta_l(x)$.

Plugging this requirement back into our discriminant function gives:

$$\vec{\mu}_k^T \mathbf{\Sigma^{-1}} \vec{x} - \frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_k + \log \pi_k = \vec{\mu}_l^T \mathbf{\Sigma^{-1}} \vec{x} - \frac{1}{2}\mu_l^T \mathbf{\Sigma^{-1}} \mu_l + \log \pi_l$$

$$(\mu_k - \mu_l)^T \mathbf{\Sigma^{-1}} x - \frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_k + \frac{1}{2}\mu_l^T \mathbf{\Sigma^{-1}} \mu_l + \log \frac{\pi_k}{\pi_l} = 0$$

$$(\mu_k - \mu_l)^T \mathbf{\Sigma^{-1}} x - \frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_k + \left(\frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_l - \frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_l\right) + \frac{1}{2}\mu_l^T \mathbf{\Sigma^{-1}} \mu_l + \log \frac{\pi_k}{\pi_l} = 0$$

$$(\mu_k - \mu_l)^T \mathbf{\Sigma^{-1}} x - \frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_k + \left(\frac{1}{2}\mu_k^T \mathbf{\Sigma^{-1}} \mu_l - \frac{1}{2}\mu_l^T \mathbf{\Sigma^{-1}} \mu_k\right) + \frac{1}{2}\mu_l^T \mathbf{\Sigma^{-1}} \mu_l + \log \frac{\pi_k}{\pi_l} = 0$$

Boundary Equation

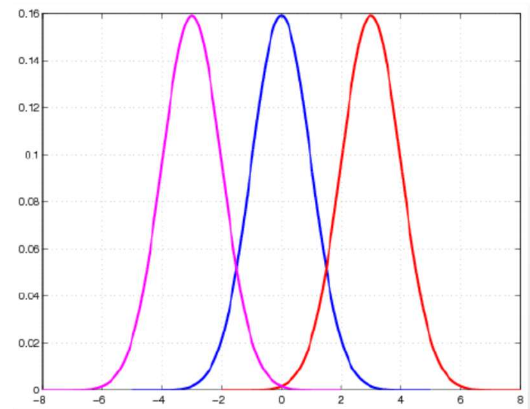
$$(\mu_k - \mu_l)^T \mathbf{\Sigma^{-1}} x - \frac{1}{2}(\mu_k + \mu_l)^T \mathbf{\Sigma^{-1}} (\mu_k - \mu_l) + \log \frac{\pi_k}{\pi_l} = 0$$

So… what does this decision boundary really mean?

Some examples:

1) 1D case: (x is one dimensional, so the mean for each class is one dimensional), as in the picture:

Suppose we have 2 classes: k=1, l=2.
Let $\pi_1=\pi_2=1/2$.
Let $\mu_1=0$ and $\mu_2=2$.
Let $\sigma=1$. (Note that in 1D, $\Sigma^{-1}=1/\sigma^2$.)



Plugging this into our decision boundary equation gives:

$$\frac{(\mu_2-\mu_1)}{\sigma^2}x - \frac{(\mu_2+\mu_1)(\mu_2-\mu_1)}{2\sigma^2} + \log\frac{\pi_2}{\pi_1} = 0$$

$$\frac{(\mu_2-\mu_1)}{\sigma^2}x = \frac{(\mu_2+\mu_1)(\mu_2-\mu_1)}{2\sigma^2} - \log\frac{\pi_2}{\pi_1}$$

$$x = \frac{(\mu_2+\mu_1)}{2} - \frac{\sigma^2}{(\mu_2-\mu_1)}\log\frac{\pi_2}{\pi_1}$$

Note the above is a *general* solution. Plugging in the values from our case: x=1!

Show Matlab examples.

2) General case:

$$(\mu_k - \mu_l)^T \mathbf{\Sigma}^{-1} x - \frac{1}{2}(\mu_k + \mu_l)^T \mathbf{\Sigma}^{-1}(\mu_k - \mu_l) + \log\frac{\pi_k}{\pi_l} = 0$$

We can rewrite this as:

$$w^T x - b = 0$$

where

$$w^T = (\mu_k - \mu_l)^T \mathbf{\Sigma}^{-1}, \text{and } b = \frac{1}{2}(\mu_k + \mu_l)^T \mathbf{\Sigma}^{-1}(\mu_k - \mu_l) + \log\frac{\pi_k}{\pi_l}$$

3) Example in 2D:
Let $\pi_1=\pi_2=1/2$.
Let $\mu_1=[0,0]^T$ and $\mu_2=[2,-2]^T$.
Let $\Sigma=[1,0;0,1]$

Then $w^T=[2,-2]$, and b=4.

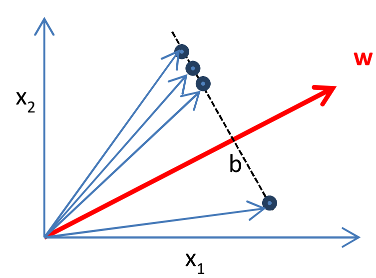
Now, of course, I can divide both sides of the boundary equation by a constant, nothing changes. Let's choose that constant to be the norm of w:

$$w^T x - b = 0 \rightarrow \frac{w^T x}{\|w\|} - \frac{b}{\|w\|} = 0$$

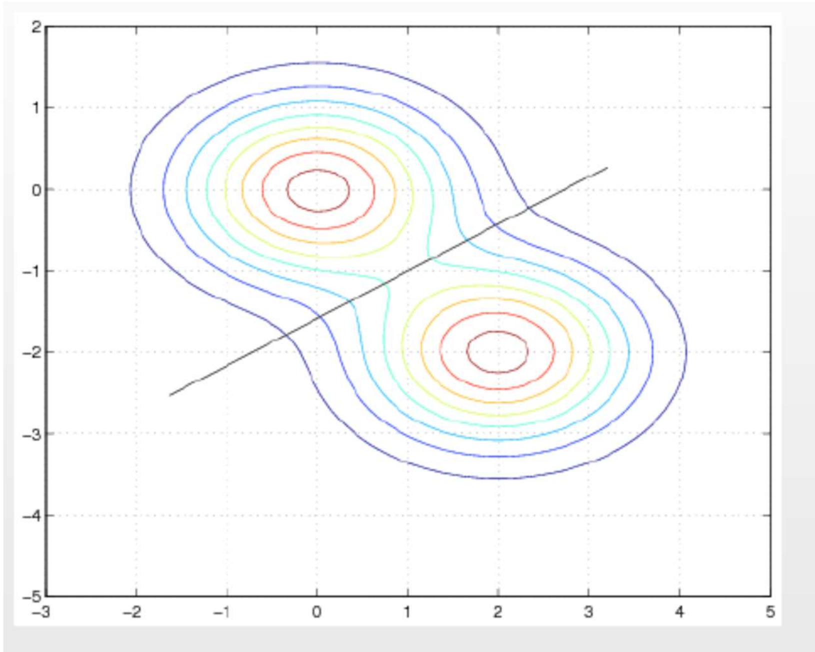
In this case, the new $w^T=[1/\sqrt{2},-1/\sqrt{2}]$, and the new $b=-4/\sqrt{8}$, or $-\sqrt{2}$.

<u>**Interpretation**</u>: the boundary is defined by the set of x's that, when you take the dot product with w, give you the value b!

In 2D, all of these x's define a line. How do I figure out where that line is?
From $w^Tx-b=0$, we have:
$w_1x_1 + w_2x_2 - b = 0$
$x_2 = b/w_2 - (w_1/w_2)*x_1$.
This is the equation for the line of the decision boundary!



Decision Boundary in 2D:



Note: in general, the classification boundary is a hyper-plane in the N-D space. We saw it's a point in 1D and a line in 2D. The x that satisfy the boundary equation

$$w^T x - b = 0$$

Make up a hyper-plane, with w the vector orthogonal to the plane.

Proof that w is orthogonal to the plane. Let x_1 and x_2 be two points that satisfy the boundary equation. Therefore:

$$w^T x_1 - b = w^T x_2 - b$$

$$w^T x_1 = w^T x_2$$

$$w^T (x_1 - x_2) = 0$$

Note that the vector x1-x2 lives in the hyper-plane, and the dot product between this vector and w is 0. Therefore, w is orthogonal to the decision hyper-plane.