Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

Linear Classifiers

If you wanted to break up your data into two parts ($+1$ and -1), how might you do it? Let's explore that question.

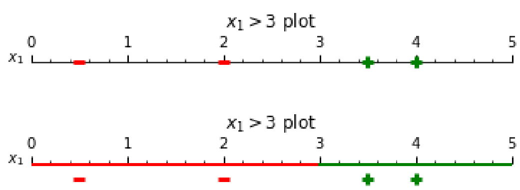
1-D Linear Classifiers

As usual, we'll start with the **simplest** case we can think of: 1-D. So, we only have one variable x_1 to **classify** with.

The simplest version might be to just **split** our space in **half**: those above or below a certain **value**. This is our parameter, C.

$$x_1 > C \qquad \text{or} \qquad x_1 - C > 0 \tag{1}$$

Example: For the below data (where green gives positive and red gives positive), could classify positive as $x_1 > 3$.



We plot everything above $x = 3$ as **positive**, and **negative** otherwise.

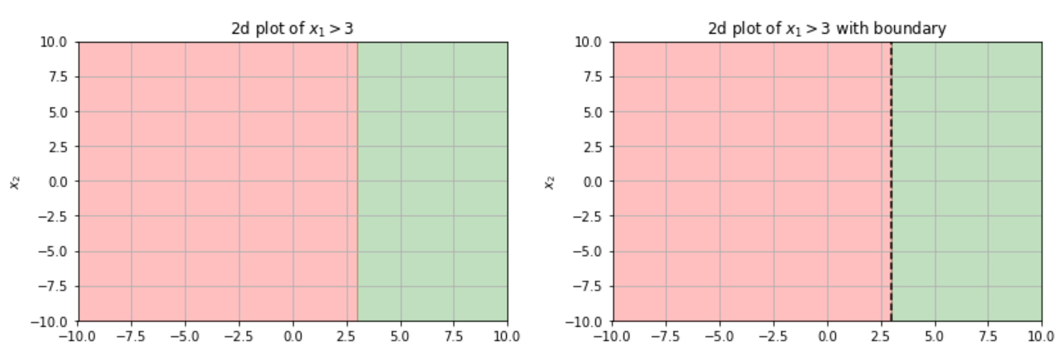
We could also call it θ_0, in the spirit of our θ notation for parameters.

$$x_1 + \theta_0 > 0 \tag{2}$$

1-D classifiers in 2-D

Let's add a variable and see how our classifier looks on a 2-D plot.

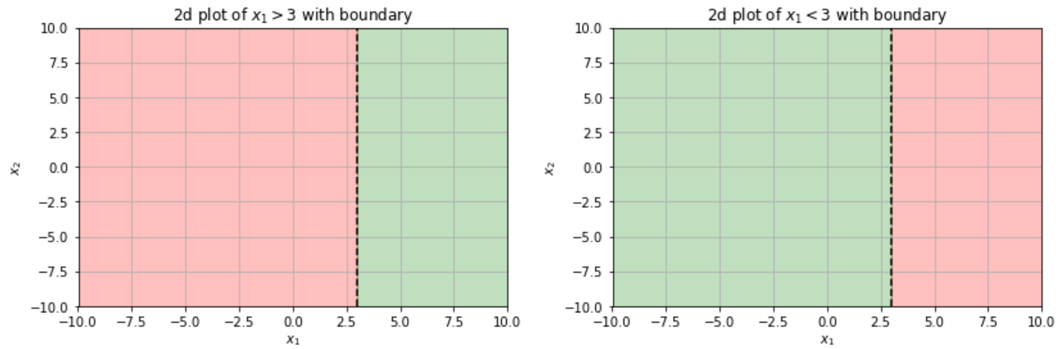
We'll omit the data points for now.



On the right, we've drawn the **dividing** line between our two regions.

Interesting - the **boundary** between positive and negative is defined by a **vertical line**.

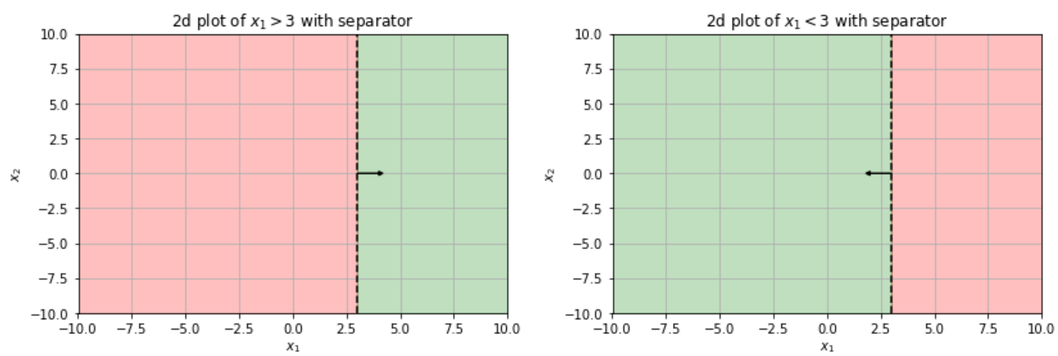
Or, almost. Compare $x_1 > 3$ and $x_1 < 3$:



These two plots have the same line, but have their sides flipped.

So, we have a **line** that gives us the boundary, but we **also** need to include information about which way is the **positive** direction.

What tool best represents **direction**? We could use angles, but we haven't used that much so far. Instead, let's use a **vector** to **point** in the right direction.



Now, it's clear which plot is which, just using our **line** and **vector**!

The object that represents our classification is called a **separator**!

Since our variables are x_1 and x_2, this is a separator in **input space**.

Definition 1

A **separator** defines how we **separate** two different classes with our **hypothesis**.

It includes

- The **boundary**: the **surface** where we **switch** from one **class** to another.
- The **orientation**: a **description** of which **side** of the boundary is assigned to **which class**.

We call it "orientation" because you could imagine "flipping over" the space, so the positive and negative regions are swapped.

For example, let's take our specific separator from above.

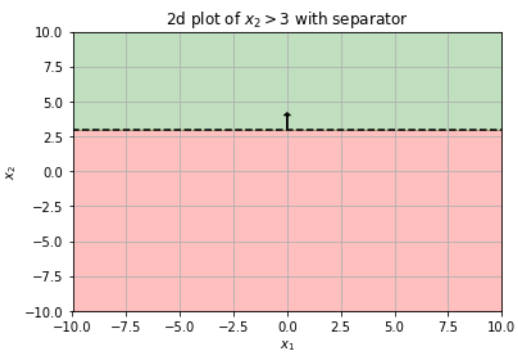
Concept 2

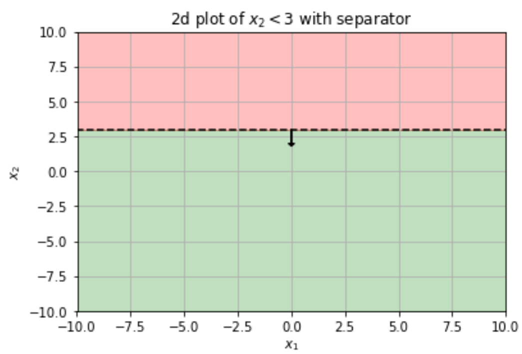
We can define our **1-D separator** using

- The **boundary** between the **positive** and **negative** regions: in 2-D input space, this looks like a vertical or horizontal **line**.
- A **vector** pointing towards whichever side is given a $+1$ **value**.

A second 1-D separator, and our problem

What if we use x_2 to **separate** our data?



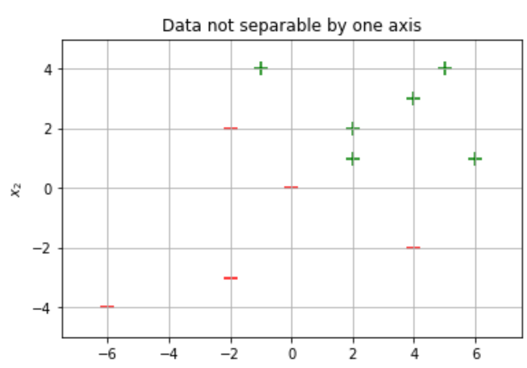


Instead of having a vertical separator, we have a **horizontal** one.

We get the same sort of plot along the **other axis**!

So, this is cool so far, but it's not a very **powerful** model: we can only handle a situation where the data is evenly divided by **one axis**.

And if that's the case, what's the point of our **other** variable?



There's no vertical or horizontal line we can use to split this space!

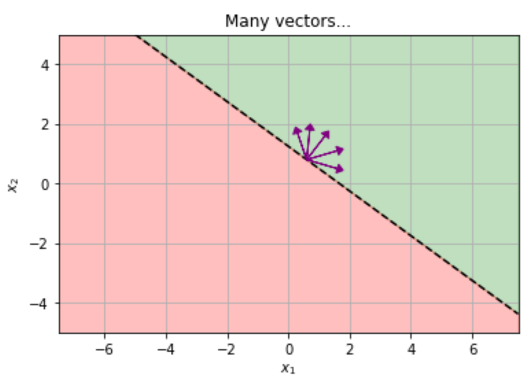
The 2-D Separator: What vector do we use?

Just looking at our example, we might wonder, "well, if we can use **vertical** lines or **horizontal** lines, can't we just use a line in **another** orientation?

It turns out, we **can**!

If we allow lines at an angle, we can classify all of our data correctly!

So, we've got our **boundary**. But we still need a vector to tell us which side is **positive**. But there are **many** possible vectors we could choose:



All of these vectors point towards the **correct** side of the plane. Is there a **best** one to use?

Above, we used the vector that was **vertical** or **horizontal**. This makes sense: if we're doing $x_1 > 3$, it seems reasonable to have the arrow **point** in the positive-x_1 direction.

But this vector also happened to be **perpendicular** to our **line**: this is the line's **normal vector**, $\hat{n}$. This vector has a couple nice properties:

- It is **unique**: in 2-D, there is only 1 **normal** direction.
- It points directly **away** from the plane.
- If our plane is at the **origin**, any point with a **positive** $\hat{n}$ component is on the **positive** side.

The opposite side is just $-\hat{n}$.

This will be important later!

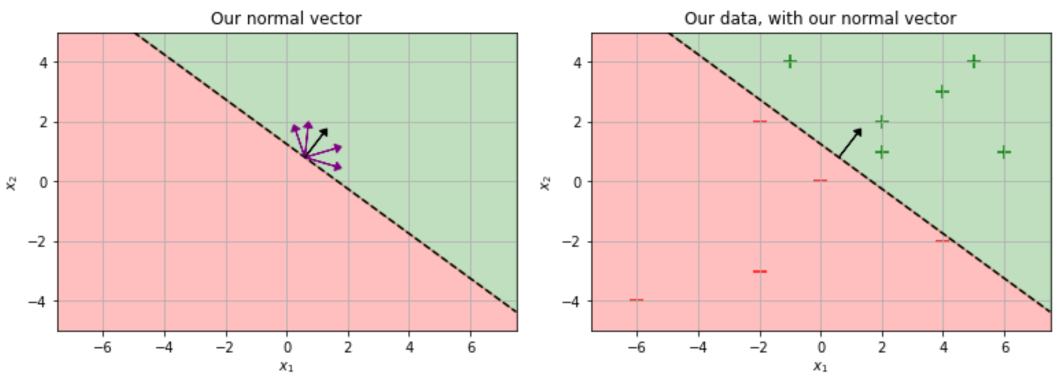
So, we have a **unique** vector that tells us which side is **positive**. Let's go with that!

Concept 3

Every **line** in 2-D has a **unique normal vector** that can be used to **define** the **angle/direction** of the line.

The **direction** the vector is "facing" is also called the **orientation**.

Our normal vector for the above separator:



We can define our plane using the **normal** vector!

It's clear that this vector in some way is a **parameter**: if we change this vector, we get a different **orientation**, and a different **classifier**.

We have **represented** parameters in the past using θ. We need **two** different θ_k: one for the x_1 component, another for the x_2 component.

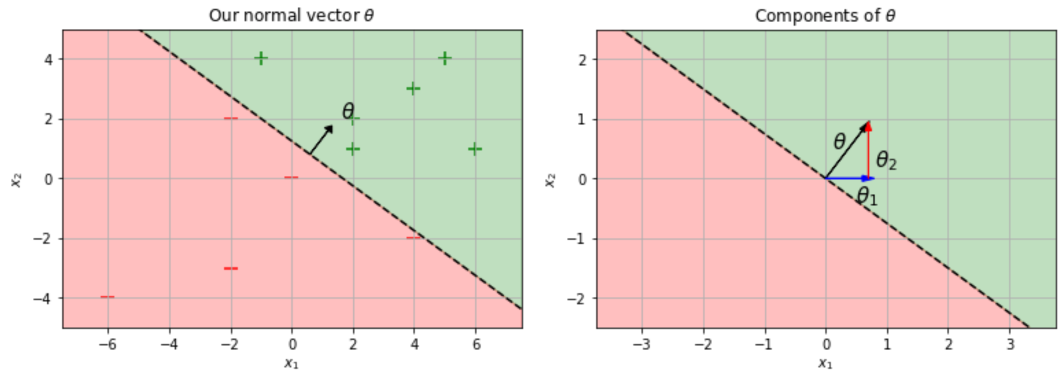
So, we'll use that.

Notation 4

The vector θ represents the **normal vector** to our line in 2D.

$$\hat{n} = \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

We add this to our diagram:

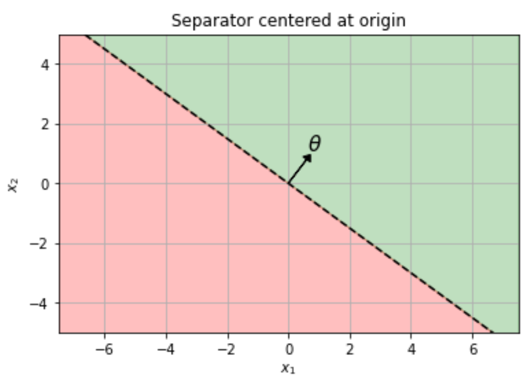


θ is our normal vector!

Nice work so far. The next question is: how do we describe this separator **mathematically**?

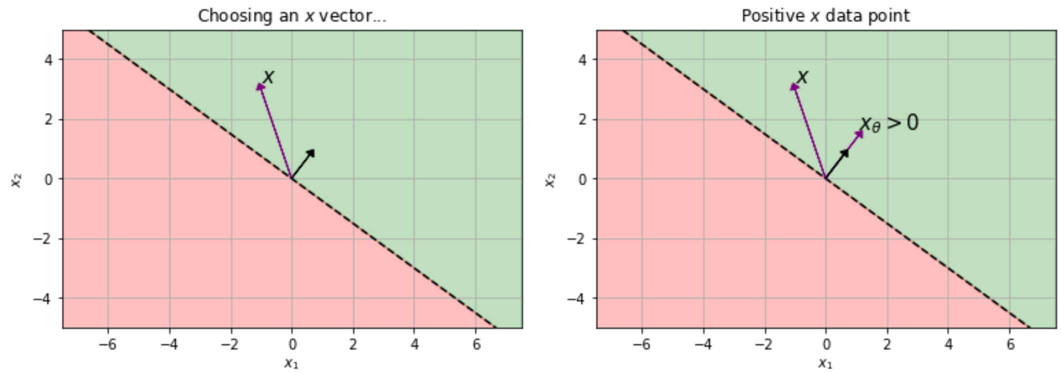
2D Separator - Matching components

As always, we'll **simplify** the problem to make it more manageable: for now, we'll assume our **separator** is centered at the **origin**.

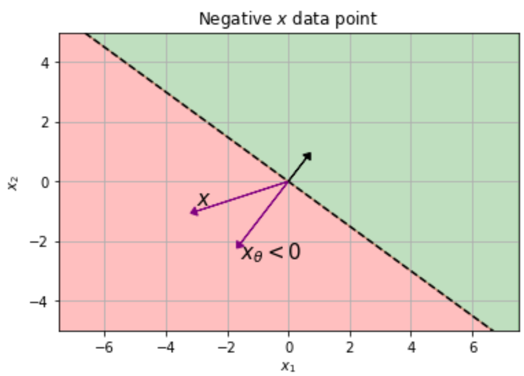


So, we have our vector, $\hat{n}$. As we mentioned above, anything on the **same** side as $\hat{n}$ is **positive**, and anything on the **opposite** side is **negative**.

For a line on the origin, "On the same side of the line" can be interpreted as "has a positive $\hat{n}$ component". We'll find that component next.



This vector has a **positive** component in the θ direction.



This vector has a **negative** component in the θ direction.

How do we represent "on the same side" mathematically? How do we **find** whether the component is **positive** or **negative**? We use the **dot product**.

The Dot Product (Review)

How to calculate the dot product should be familiar to you, but we'll talk about some **intuition** that you may not be exposed to.

Concept 5

You can use the **dot product** between unit vectors to measure their "**similarity**": if two vectors are more **similar**, they have a **larger** dot product.

In the most clear cases, take unit vectors $\hat{a}$ and $\hat{b}$:

- If they are in the **exact same** direction, $\hat{a} \cdot \hat{b} = 1$
- If they are in the **exact opposite** direction, $\hat{a} \cdot \hat{b} = -1$
- If they are **perpendicular** to each other, $\hat{a} \cdot \hat{b} = 0$

Remember, **unit vectors** have a length of 1.

What about non-unit vectors?

These unit vectors are then scaled up by the **magnitude** of each of our vectors. Because magnitudes are **always positive**, the dot product sign doesn't change.

Concept 6

You can use the **dot product** between non-unit vectors to measure their "similarity" **scaled by their magnitude**.

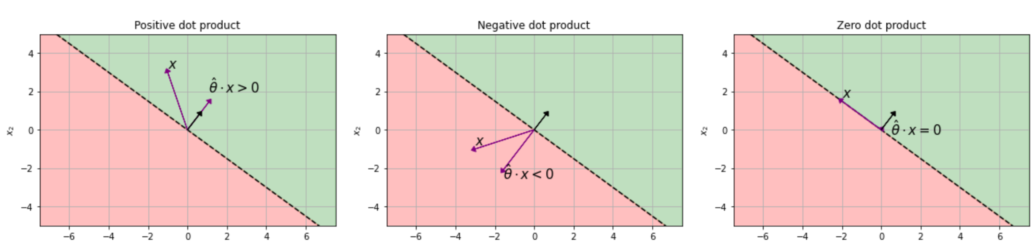
If two vectors are more **similar**, they have a **larger** dot product.

- If the vectors are **less** than 90° apart, they are more similar: they will share a **positive** component: $\vec{a} \cdot \vec{b} > 0$
- If the vectors are **more** than 90° apart, they will share a **negative** component: $\vec{a} \cdot \vec{b} < 0$
- If they are **perpendicular** (90°) to each other, $\vec{a} \cdot \vec{b} = 0$

Using the dot product

So, the **sign** of the dot product is a useful tool. If a point is on the line, it is **perpendicular** to θ, our **normal vector**.

So, if a point has a **positive** dot product, it is on the **same side** as θ, and if it's **negative**, it's on the opposite side.



Our various dot products can show us where in the space we are.

So, we can classify things based on the **sign** of it. Written as an equation, we can define the sign function:

Key Equation 7

For a **linear separator** centered on the **origin**, we can do **binary classification** using the hypothesis

$$h(x;\theta) = \text{sign}(\theta \cdot x) = \begin{cases} +1 & \text{if } \theta \cdot x > 0 \\ -1 & \text{otherwise} \end{cases}$$