Support Vector Machines
Part 2:
The Polynomial Kernel

Specifically, we're going to talk about the **Polynomial Kernel's** parameters…

Polynomial Kernel

$$\boxed{(a \times b + r)^d}$$

…and how the **Polynomial Kernel** calculates high-dimensional relationships

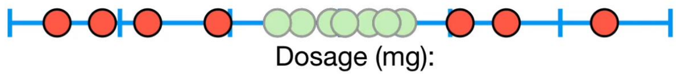
Polynomial Kernel

$$(a \times b + r)^d$$

$$\begin{aligned}
(a \times b + \frac{1}{2})^2 &= (a \times b + \frac{1}{2})(a \times b + \frac{1}{2}) \\
&= ab + a^2b^2 + \frac{1}{4} \\
&= (a, a^2, \frac{1}{2}) \cdot (b, b^2, \frac{1}{2})
\end{aligned}$$

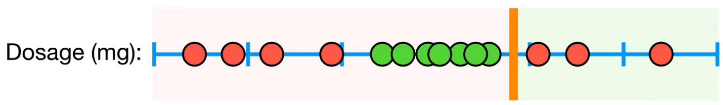
In the **StatQuest** on **Support Vector Machines**, we had a **Training Dataset** based on **Drug Dosages** measured in a bunch of patients.

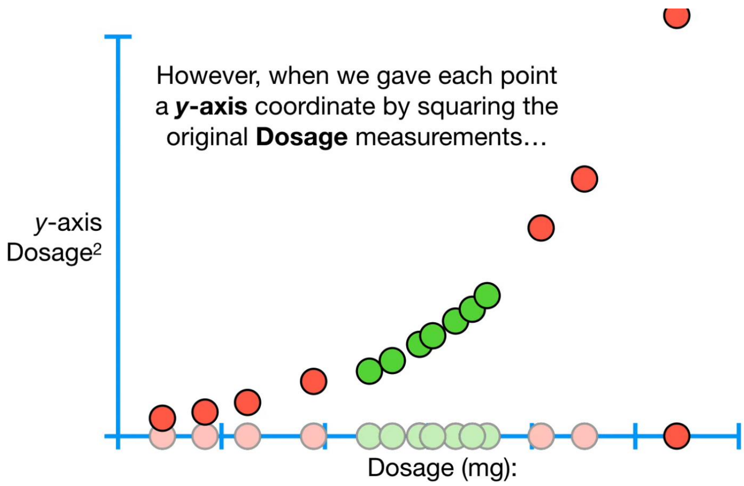
The **red dots** represented patients that were ***not cured***…

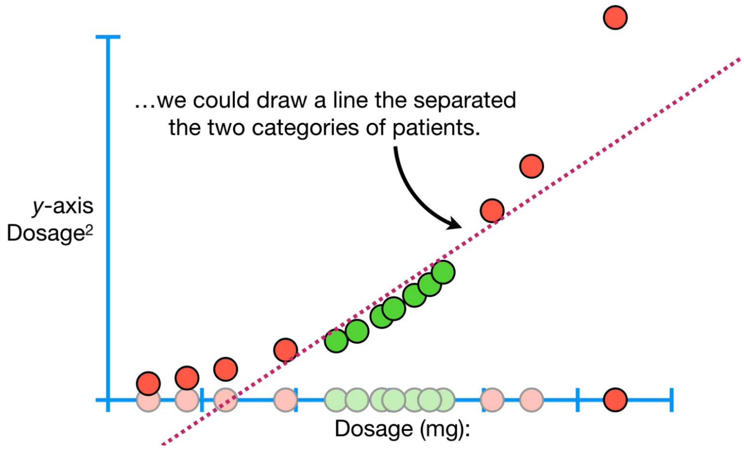


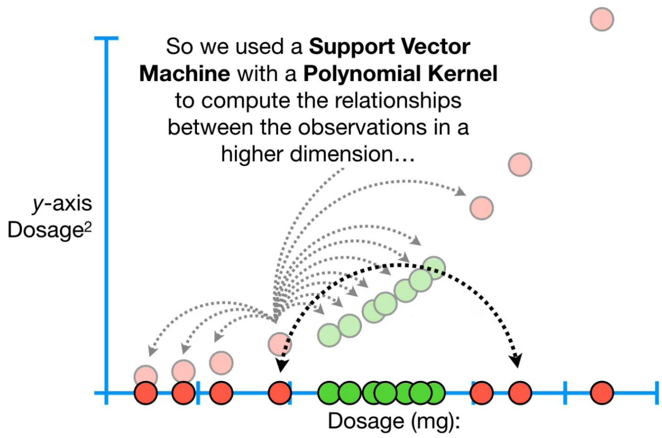
…and the **green dots** represented patients that were ***cured***.

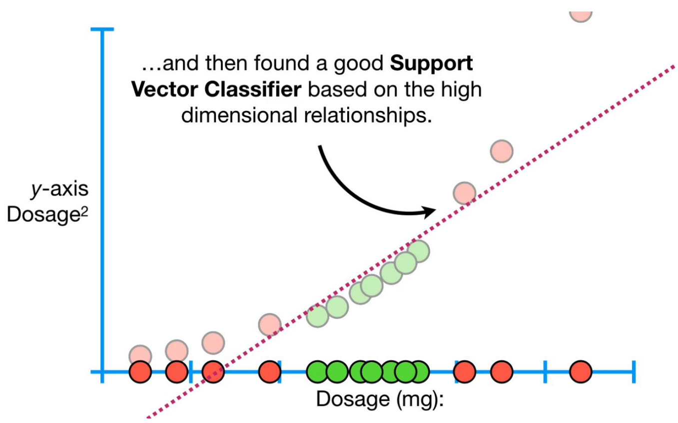
Because this **Training Dataset** had so much overlap, we were unable to find a satisfying **Support Vector Classifier** to separate the patients that were cured from the patients that were not cured.




However, when we gave each point a **y-axis** coordinate by squaring the original **Dosage** measurements…
y-axis
Dosage²
Dosage (mg):


…we could draw a line the separated the two categories of patients.
y-axis
Dosage²
Dosage (mg):


So we used a **Support Vector Machine** with a **Polynomial Kernel** to compute the relationships between the observations in a higher dimension…
y-axis
Dosage²
Dosage (mg):

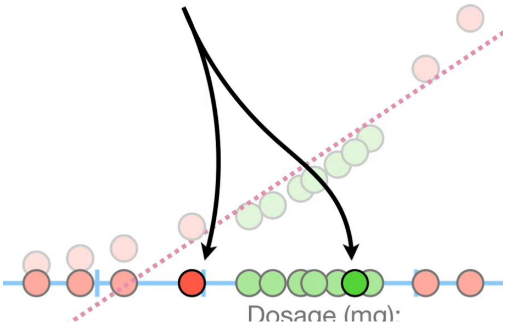

…and then found a good **Support Vector Classifier** based on the high dimensional relationships.
y-axis
Dosage²
Dosage (mg):

$$\boxed{(a \times b + r)^d}$$

The **Polynomial Kernel** that I used looks like this.

$$(\boxed{a \times b} + r)^d$$

a and ***b*** refer to two different observations in the dataset.



$$(a \times b + \boxed{r})^d$$

…***r*** determines the coefficient of the polynomial…

$$(a \times b + r)^{\boxed{d}}$$

…and, like I mentioned in the earlier **StatQuest**, ***d*** sets the degree of the polynomial.

$$(a \times b + \frac{1}{2})^2$$

In my example, I set ***r*** = **1/2** and ***d*** = **2**.

$$(a \times b + \frac{1}{2})^2 = (a \times b + \frac{1}{2})(a \times b + \frac{1}{2})$$

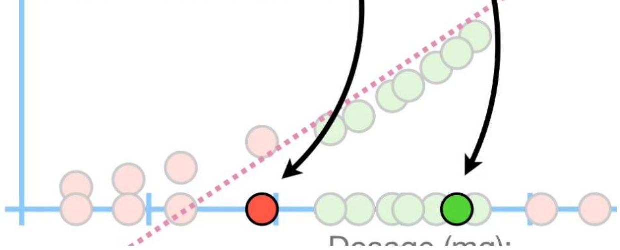
Since we are squaring the term, we can expand it to be the product of two terms.

$$\begin{aligned}(a \times b + \frac{1}{2})^2 &= (a \times b + \frac{1}{2})(a \times b + \frac{1}{2}) \\ &= ab + a^2b^2 + \frac{1}{4} \\ &= (a, a^2, \frac{1}{2}) \cdot (b, b^2, \frac{1}{2})\end{aligned}$$

...is equal to this **Dot Product**.

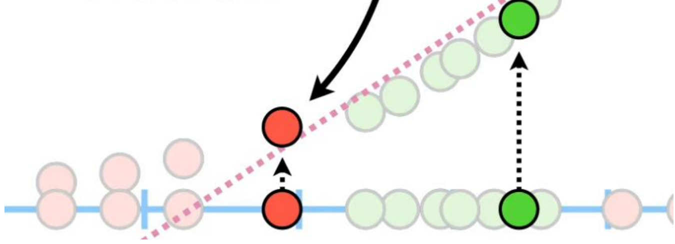
$$= (\boxed{a,} a^2, \frac{1}{2}) \cdot (\boxed{b,} b^2, \frac{1}{2})$$

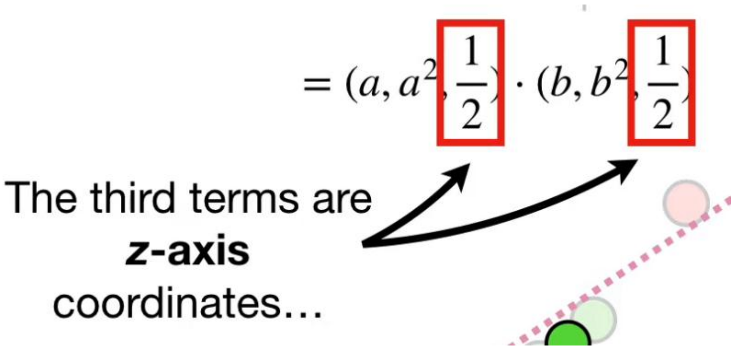
The first terms are the ***x*-axis** coordinates...



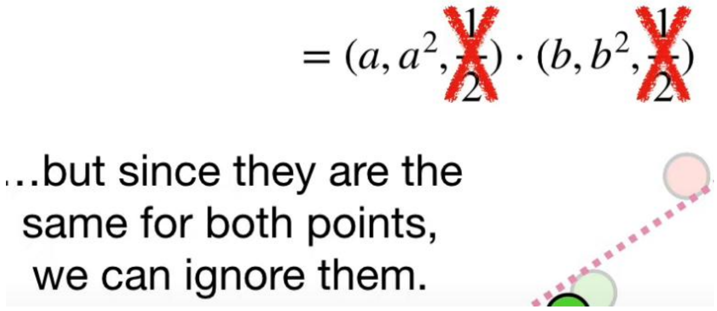
$$= (a, \boxed{a^2,} \frac{1}{2}) \cdot (b, \boxed{b^2,} \frac{1}{2})$$

...and the second terms are the ***y*-axis** coordinates.

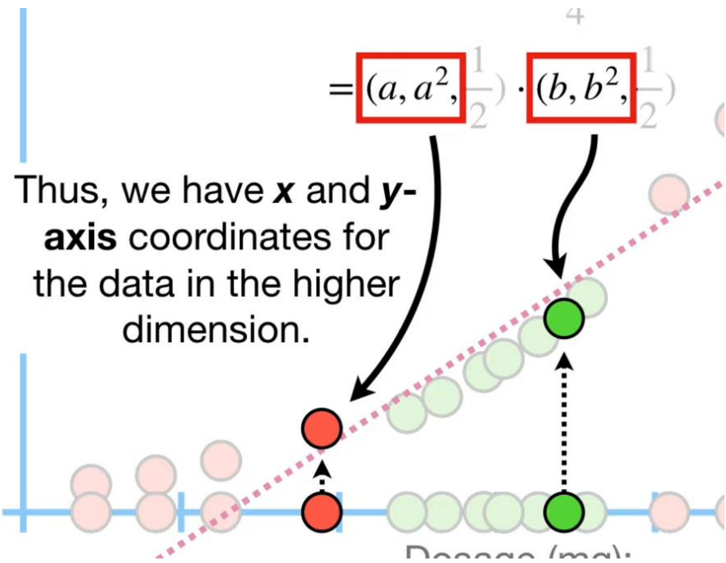




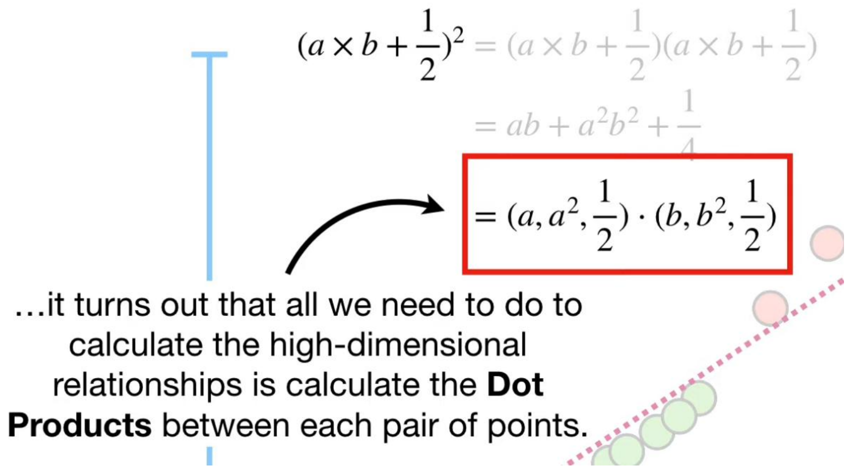
$= (a, a^2, \boxed{\frac{1}{2}}) \cdot (b, b^2, \boxed{\frac{1}{2}})$
The third terms are **z-axis** coordinates…



$= (a, a^2, \cancel{\frac{1}{2}}) \cdot (b, b^2, \cancel{\frac{1}{2}})$
…but since they are the same for both points, we can ignore them.



$= \boxed{(a, a^2,} \frac{1}{2}) \cdot \boxed{(b, b^2,} \frac{1}{2})$
Thus, we have ***x*** and ***y*-axis** coordinates for the data in the higher dimension.

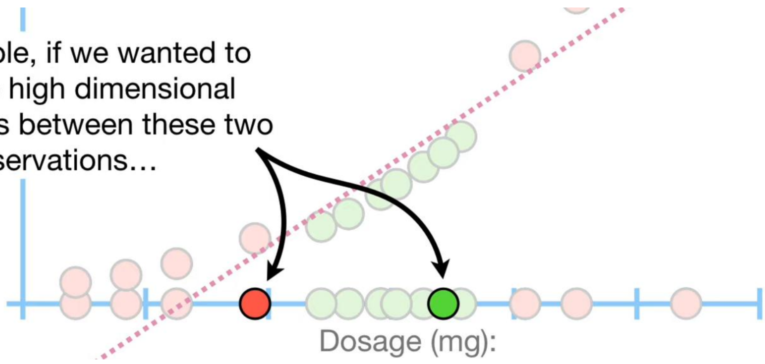


$(a \times b + \frac{1}{2})^2 = (a \times b + \frac{1}{2})(a \times b + \frac{1}{2})$
$= ab + a^2b^2 + \frac{1}{4}$
$\boxed{= (a, a^2, \frac{1}{2}) \cdot (b, b^2, \frac{1}{2})}$
…it turns out that all we need to do to calculate the high-dimensional relationships is calculate the **Dot Products** between each pair of points.

$$(a \times b + \frac{1}{2})^2$$

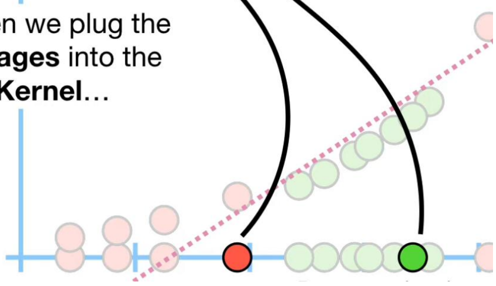
…all we need to do is plug values into the **Kernel** to get the high-dimensional relationships.

For example, if we wanted to know the high dimensional relationships between these two observations…



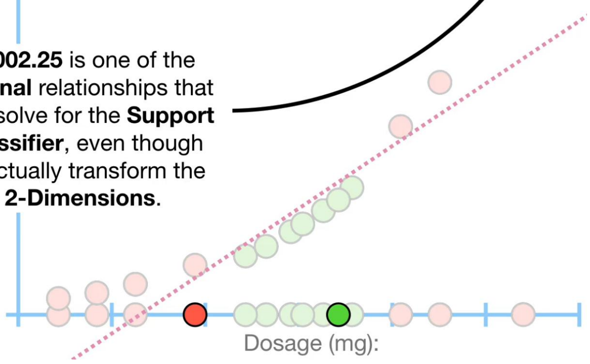
$$(9 \times 14 + \frac{1}{2})^2$$

…then we plug the **Dosages** into the **Kernel**…



$$(9 \times 14 + \frac{1}{2})^2 = (126 + \frac{1}{2})^2 = 126.5^2 = \boxed{16{,}002.25}$$

…and **16,002.25** is one of the **2-Dimensional** relationships that we need to solve for the **Support Vector Classifier**, even though we didn't actually transform the data to **2-Dimensions**.

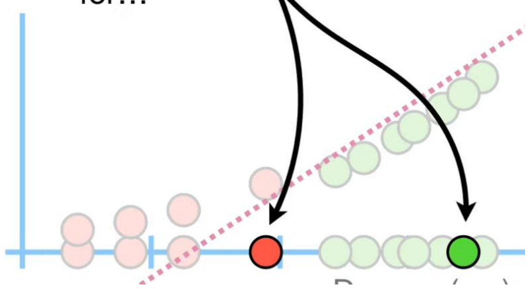


$$(a \times b + r)^d$$

To review, the **Polynomial Kernel** computes relationships between pairs of observations.

$$(\boxed{a \times b} + r)^d$$

a and ***b*** refer to the two observations we want to calculate the high dimensional relationship for…



$$(a \times b \boxed{+\, r})^d$$

…***r*** determines the polynomial's coefficient…

$$(a \times b + r)^{\boxed{d}}$$

…and ***d*** determines the degree of the polynomial.

$$(a \times b \boxed{+ r)^d}$$

NOTE: ***r*** and ***d*** are determined using **Cross Validation**.