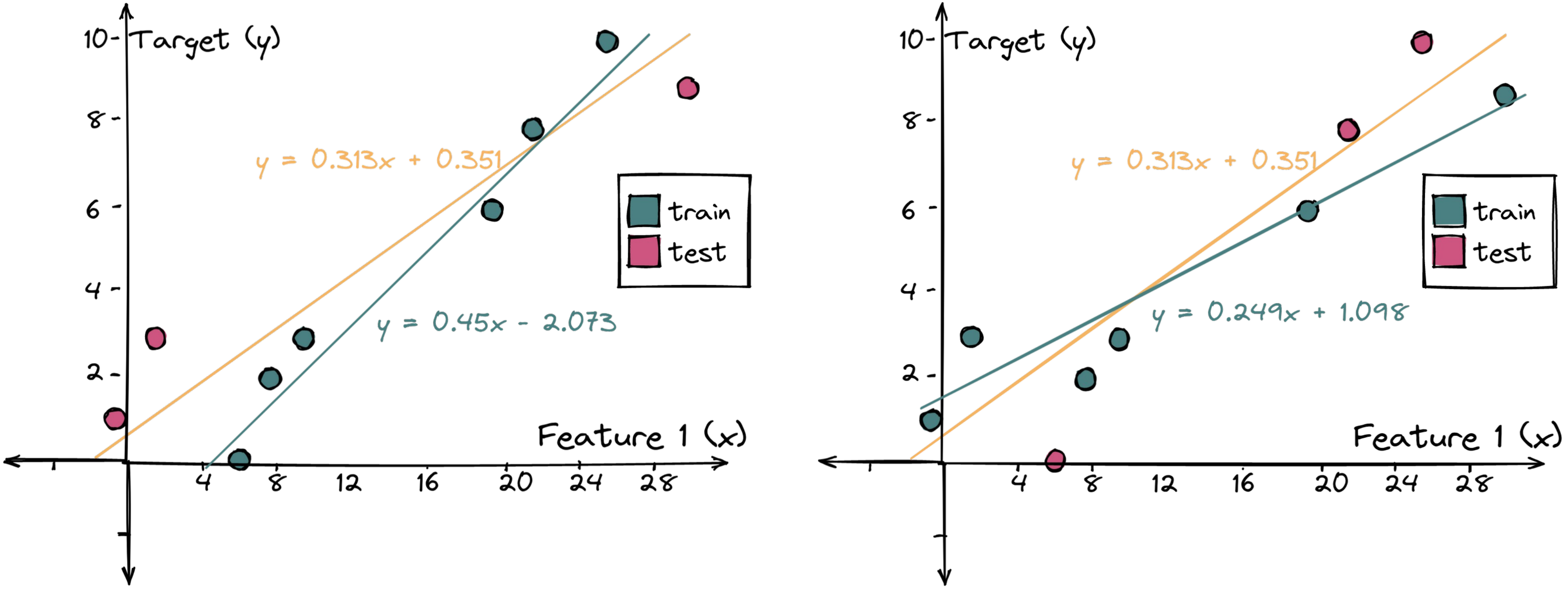
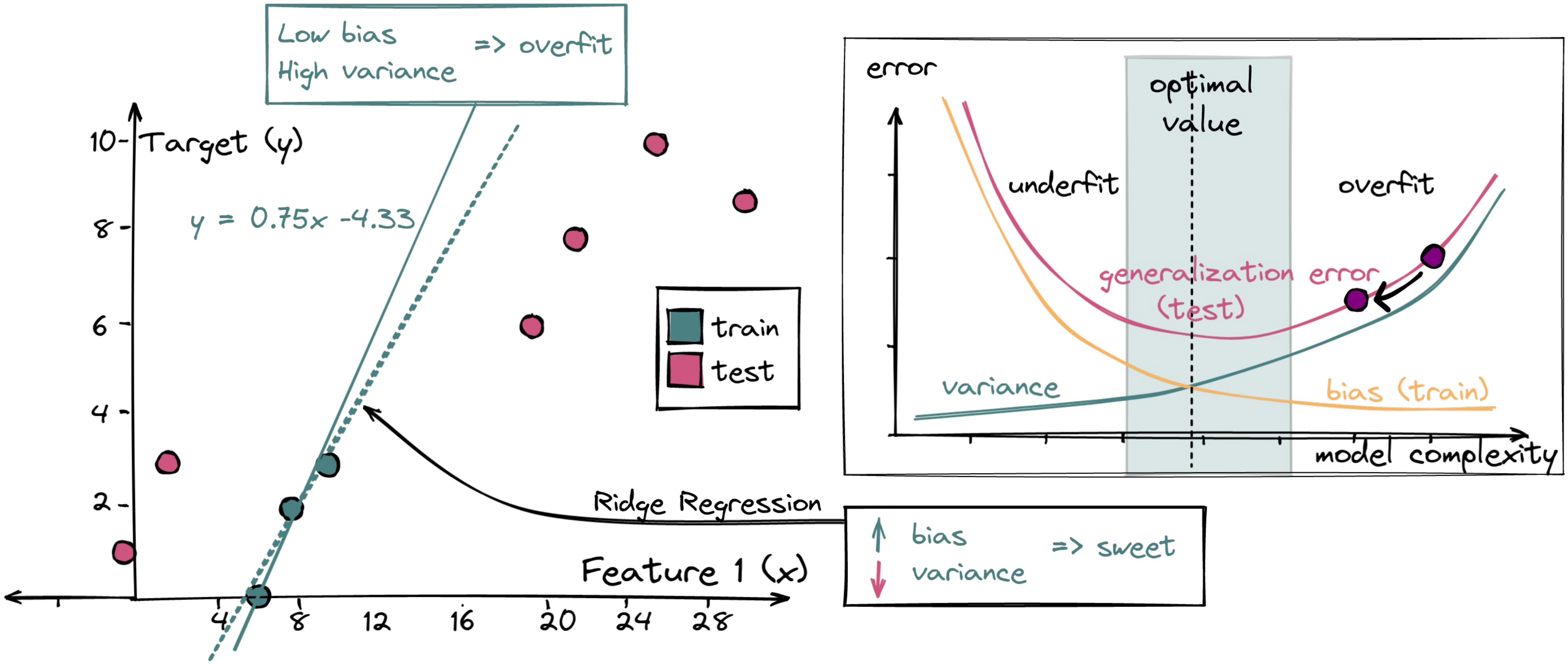
# High Variance Example



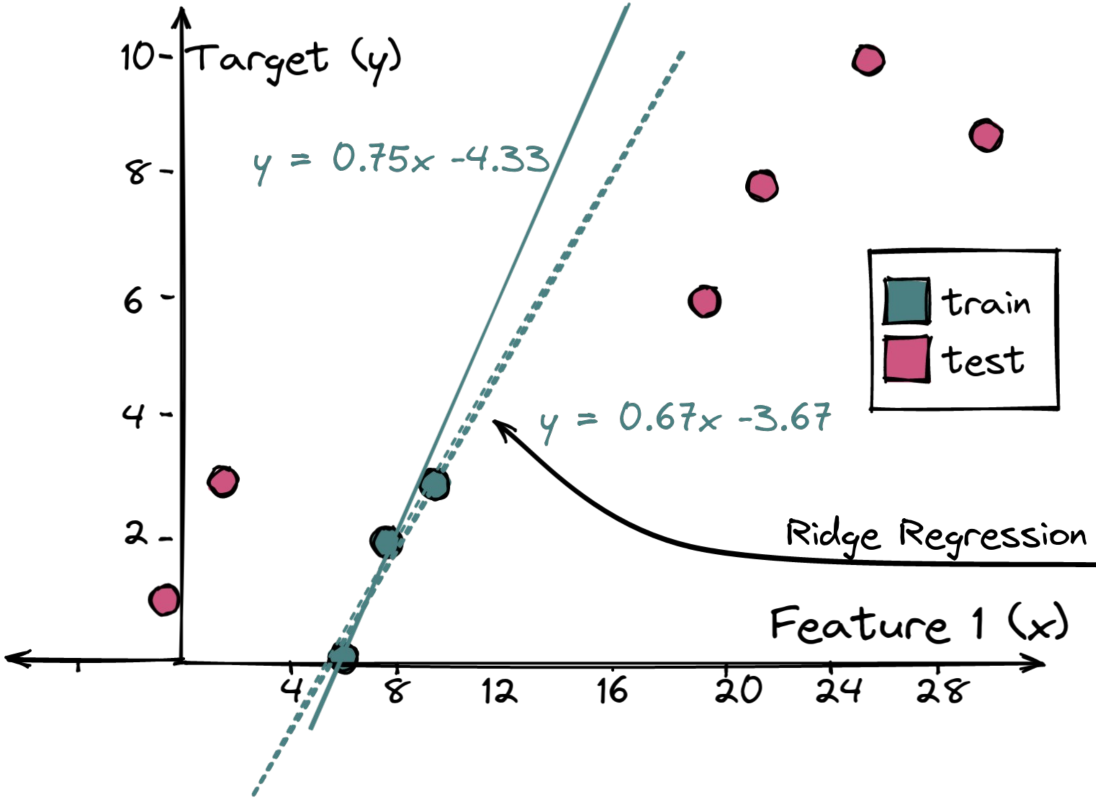

**Purpose of both Ridge and Lasso Regularization is to make y less sensible to x**

# Regularization: Ridge



# Ridge (L2) Regularization

minimizing sum of squared residuals




$$y = ax + b$$

$$L2 = sum\_sq\_residuals + \lambda * a^2$$

(assume $\lambda = 1$):

$$L2(orig) = (-0.167)^2 + 0.33^2 + (-0.167)^2 + 1 * 0.75^2 = \boxed{0.727178}$$

$$L2(new) = 0.33^2 + (-0.33)^2 + 0^2 + 1 * 0.67^2 = \boxed{0.6667}$$

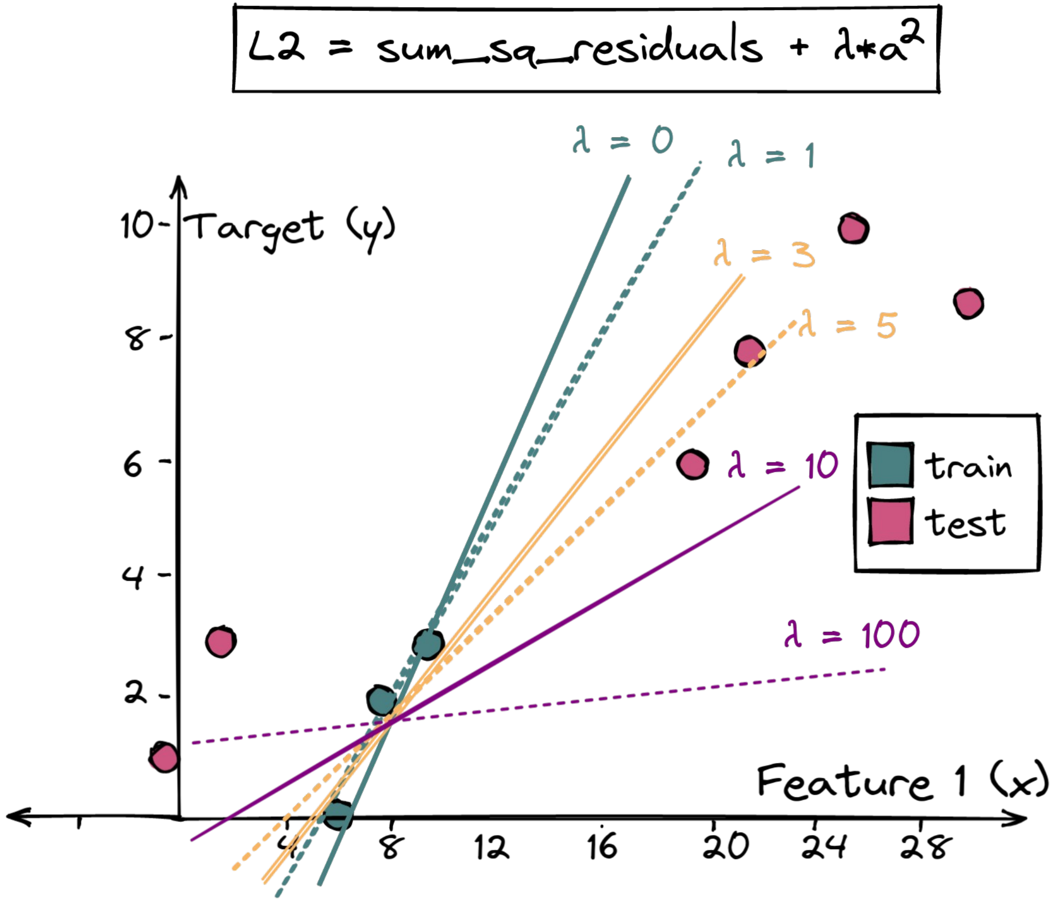
minimizing (sum of squared residuals + $\lambda * a^2$)
$\lambda$ - lambda

# Ridge (L2) Regularization

Depending on the value of λ (lambda), the linear fit will vary.

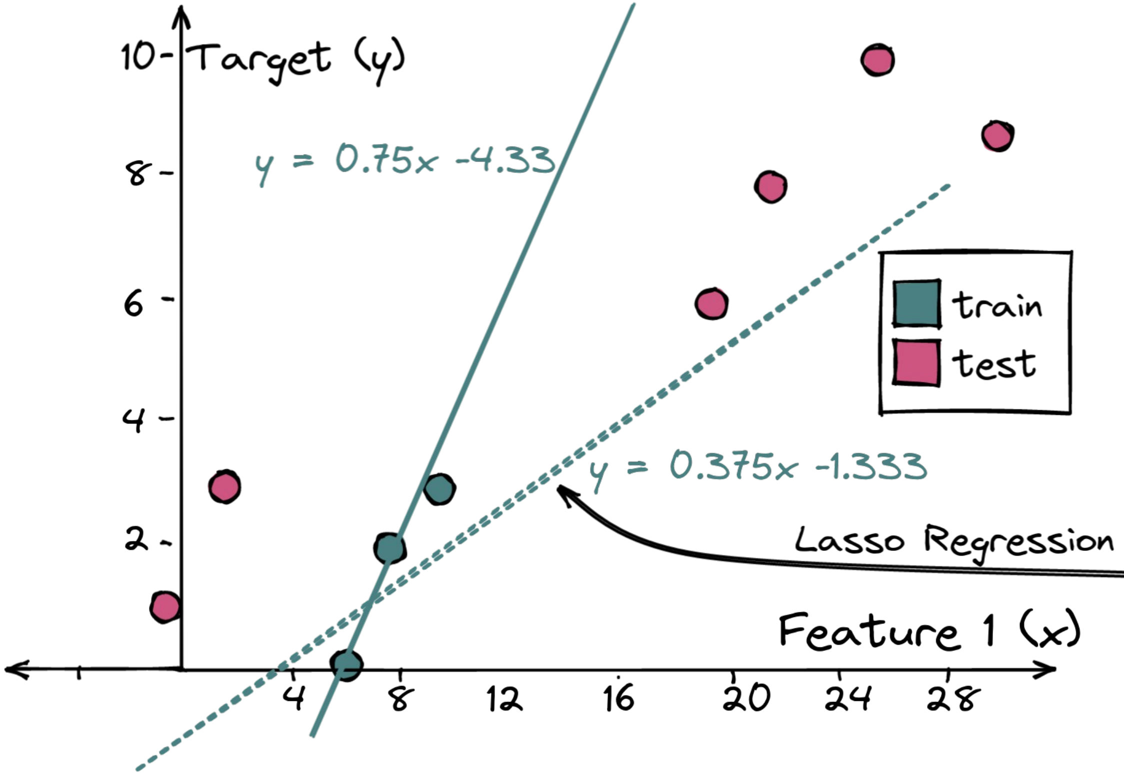
It's quite impressive, that some of the fitting lines (λ=3,5) describe unseen test data quite well

How do you think you can select λ on practice?



# Lasso (L1) Regularization

minimizing sum of squared residuals




y = ax + b

L1 = sum_sq_residuals + |λ*a|

minimizing (sum of squared residuals + |λ*a|)
λ - lambda

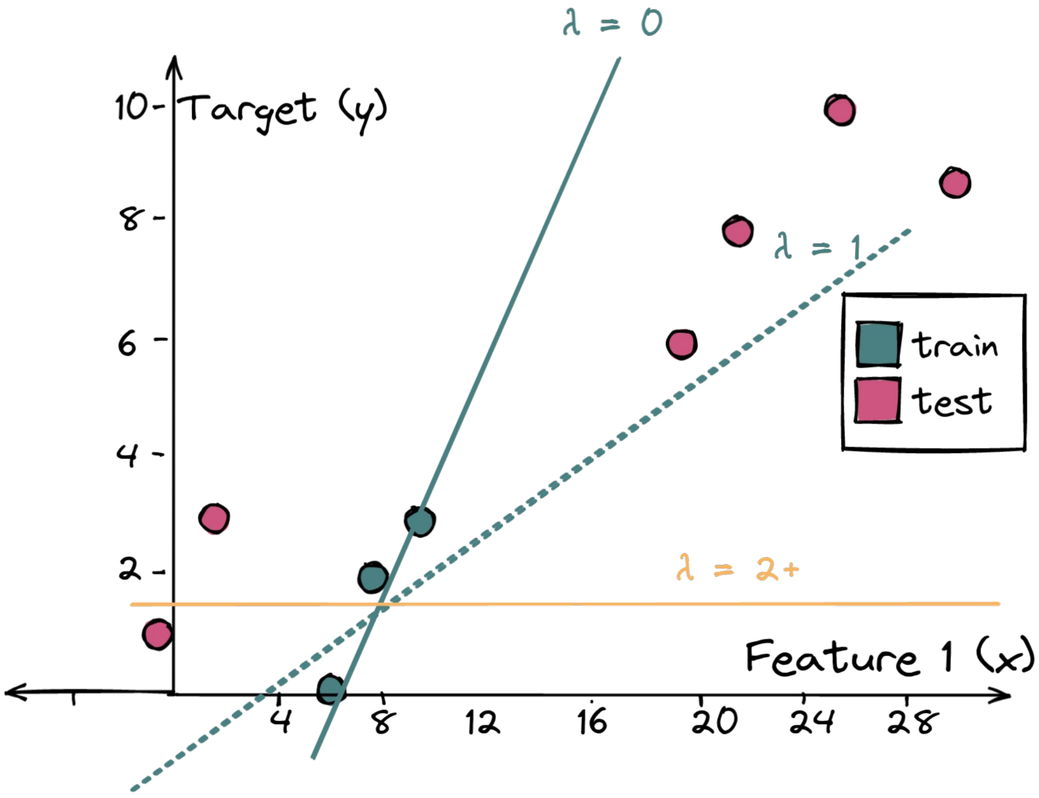
# Lasso (L1) Regularization

$$L1 = sum\_sq\_residuals + |\lambda * a|$$

Depending on the value of λ (lambda), the linear fit will also vary similar to Ridge.

Compared to Ridge, the fitting line decrease it's slope much faster

Lasso can actually represent target as constant, whereas Ridge does that asymptotically.



# Regularization Fuzz

**Multiple linear regression:**

$$\boxed{y = a_1 x + a_2 z + a_3 u + b}$$

$$L1 = \text{sum\_sq\_residuals} + \lambda * (|a_1| + |a_2| + |a_3| + \ldots)$$

$$L2 = \text{sum\_sq\_residuals} + \lambda * (a_1^2 + a_2^2 + a_3^2 + \ldots)$$

# Regularization Fuzz

**Dubai Taxi example:**

$$\boxed{y = a_1 x + a_2 z + a_3 u + b}$$

Average waiting time = intercept + $a_1$*(place) + $a_2$*(time of day) +
+ $a_3$*(day of week) + $a_4$*(your astrological animal) +
+ $a_5$*(gulf water temperature)

$$L1 = \lambda * (|a_1| + |a_2| + |a_3| + \overset{=0}{\cancel{|a_4|}} + \overset{=0}{\cancel{|a_5|}})$$

Great to use when some variables are useless

$$L2 = \lambda * (a_1^2 + a_2^2 + a_3^2 + \overset{\simeq 0}{a_4^2} + \overset{\simeq 0}{a_5^2})$$

Great to use when most variables are useful