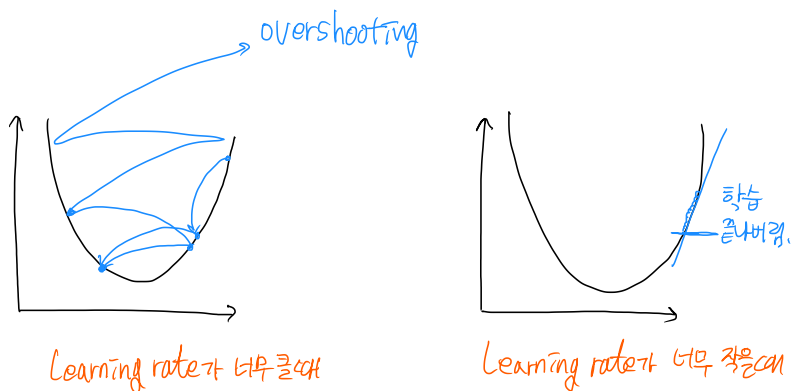


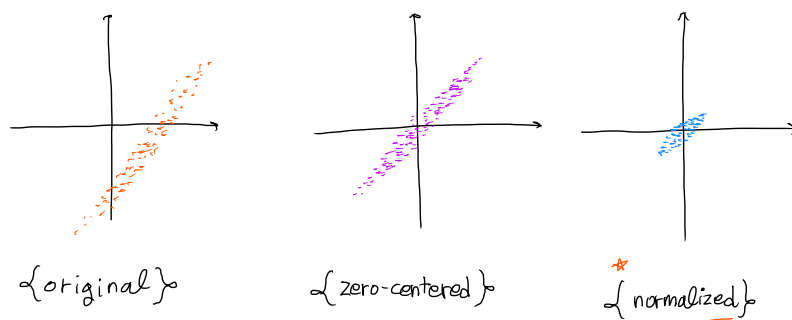
Learning rate

- 경사면을 따라서 한 발자국씩 걷는걸 학습의 과정이라 볼 때 **한 발자국의 크기로** 보면 됨
- Learning rate 가 너무 크면 값의 바깥으로 튕겨 나갈 수 있다. → **Over shooting** 이라고 한다.
- Learning rate 가 너무 작으면 step 이 끝나도 cost 의 값이 달라지지 않을 수 있다
- cost 를 출력하며 **적절한** Learning rate 의 값을 찾아야 함!





Data preprocessing

- 데이터 값의 범위가 너무 크면 역시 값이 튕겨 나갈 수 있다. 그래서 **값의 범위를** 수정해주는 작업을 하게 된다.



$$x'_j = \frac{x_j - \mu_j}{\sigma_j}$$

Overfitting

- 일반적인 Model →
- 
- overfitting Model →
- 

- # Regularization

- $$C = \frac{1}{n} \sum_i D(S(Wx^{(i)} + b), Y^{(i)}) + \boxed{\lambda \sum W^2}$$
- ↪ regularization strength

code \Rightarrow reg = 0.001 * tf.reduce_sum(tf.square(W))

0 \rightarrow 사용 x
1 \rightarrow 무조건 퍼주세요!
0 ~ 1 사이의 값을 넣는다 ~

Training / Testing Data set

- 학습데이터를 일정 부분 잘라서 테스트 용도로 사용하는 것을 **Testing Data set** 이라 한다.
- 보통 Training 70%, Testing 30%의 비율로 사용함!
- Testing Data set 으로 절대 학습시키면 안됨!

