



deeplearning.ai

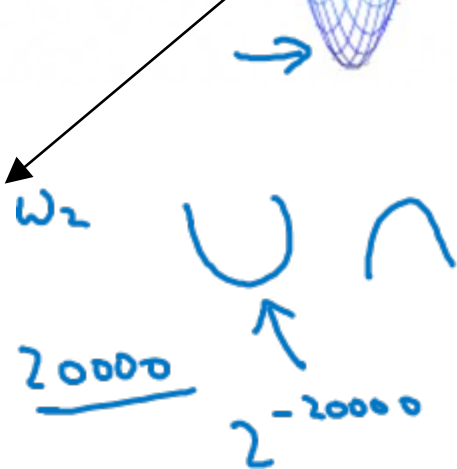
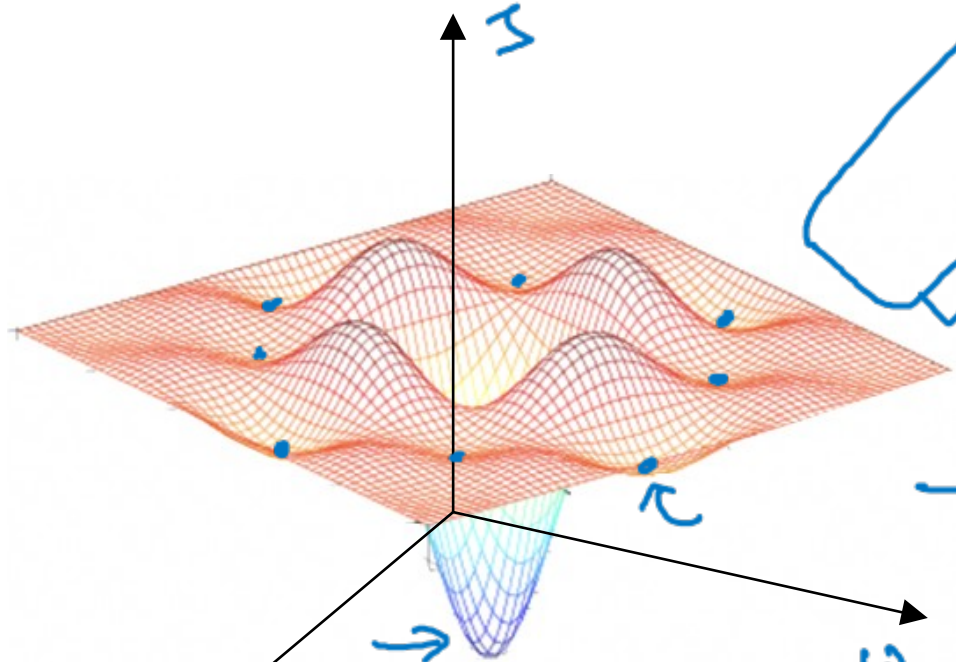
Optimization Algorithms

The problem of local optima

딥러닝 초창기에는 local minimum에
대해서 걱정을 많이 했었는데 연구가
진행되면서 관점이 많이 바뀌었다.

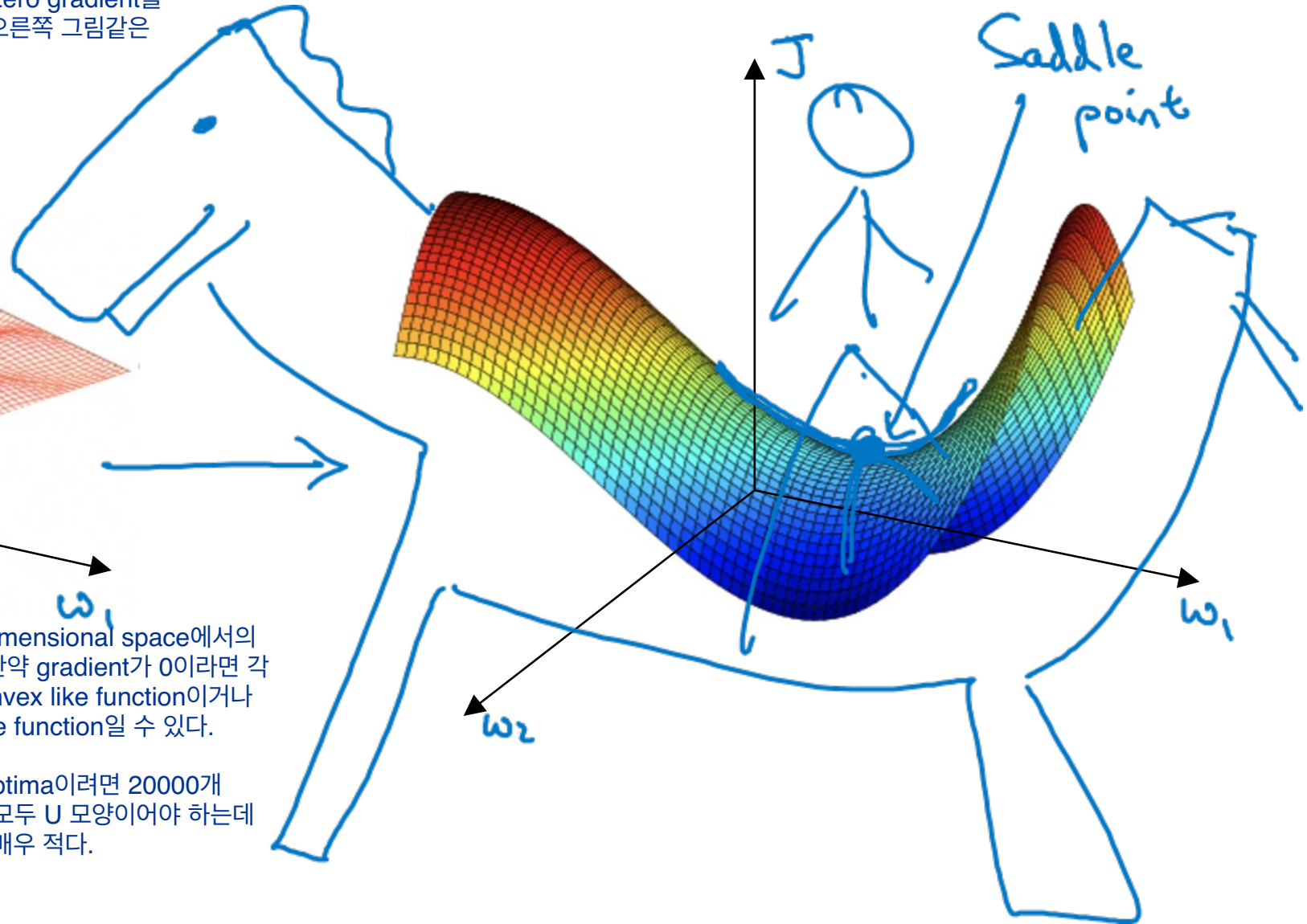
Local optima in neural networks

연구해봤더니 very high dimensional space에서는 zero gradient를 가지는 지점은 왼쪽 그림같은 local optima가 아니라 오른쪽 그림같은 saddle point 라는게 밝혀졌다.



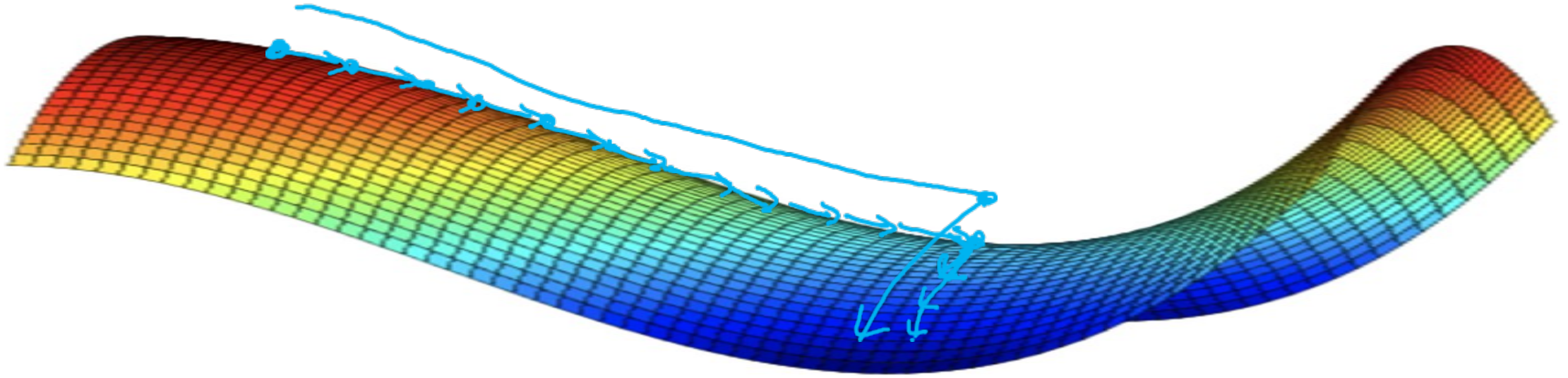
very high dimensional space에서의 function은 만약 gradient가 0이라면 각 방향으로 convex like function이거나 concave like function일 수 있다.

근데 local optima이라면 20000개 direction이 모두 U 모양이어야 하는데 그럴 확률은 매우 적다.



Problem of plateaus

그럼 local optima가 문제가 아니면 뭐가 문제냐?



결론!

- Unlikely to get stuck in a bad local optima
- Plateaus can make learning slow