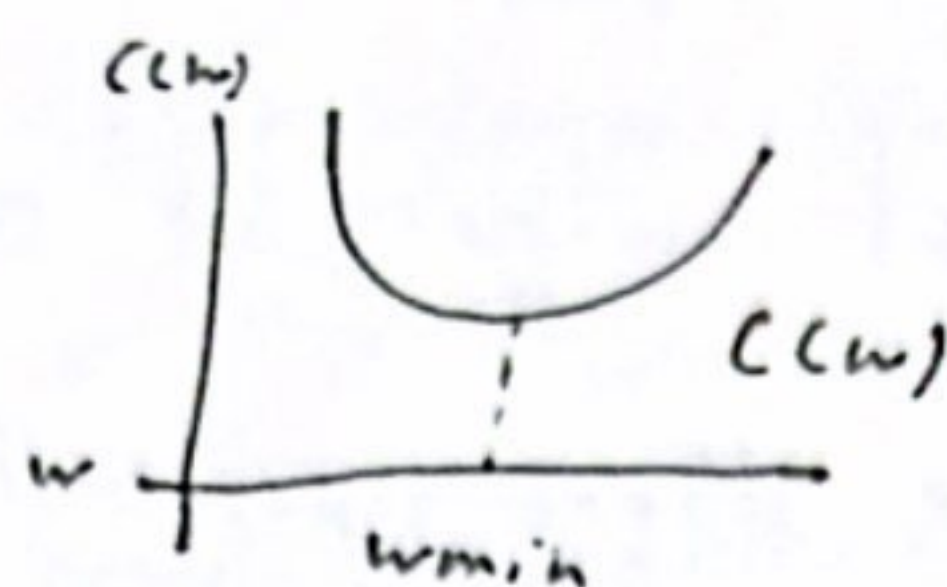


NN cost Gradient decent (12)

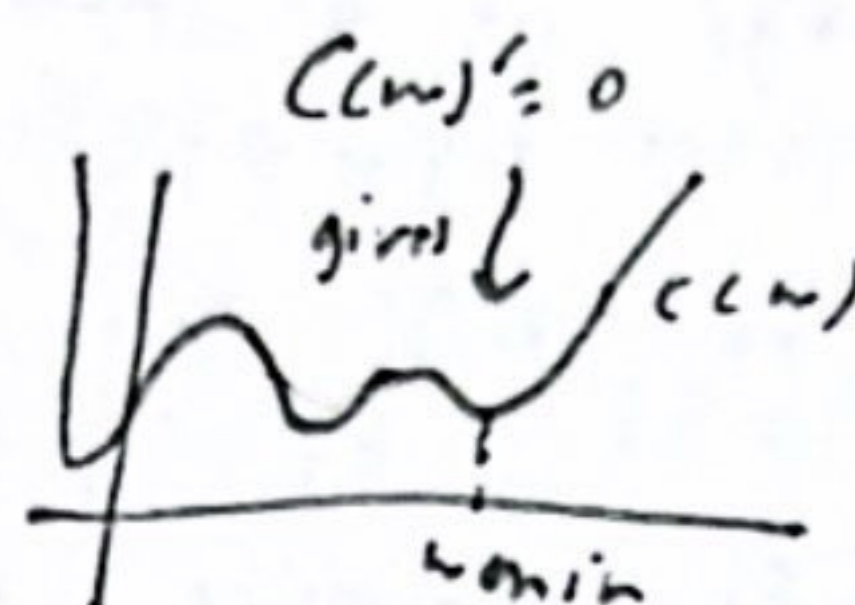
- to simplify this imagine the cost function
imagine a cost function with 2 input and 2 output

- how do you find the minimum of this function?



minimum is easy here
 $C(w)' = 0$

- but what about complicated function cannot always find min so easily



in this case it's not that simple

- and definitely not for a 13002 input function.

- a good approach is to start at any random point and figure out which direction to step in to make output less
if the slope of the function at the point is neg go left else right



- at some points you will reach the local min of function
- we cannot find the global min it's too hard but local min is good enough
- in higher dimensions it's the same concept but we use gradients to find the downhill direction so we can reach a local min.
- gradient of function gives the direction of steepest increase so we take neg of that so we can decrease function $-\nabla C$ (exy.)
the length of gradient vector is how steep it is so we have both direction and magnitude

so all we do is find gradient \rightarrow take small step in $-\nabla C$ direction
this process is ∇C Gradient decent