Embedding — Basics

- reduce dimensionality
- learn interesting traits about the text

sentence 1
sentence 2
⋮

text
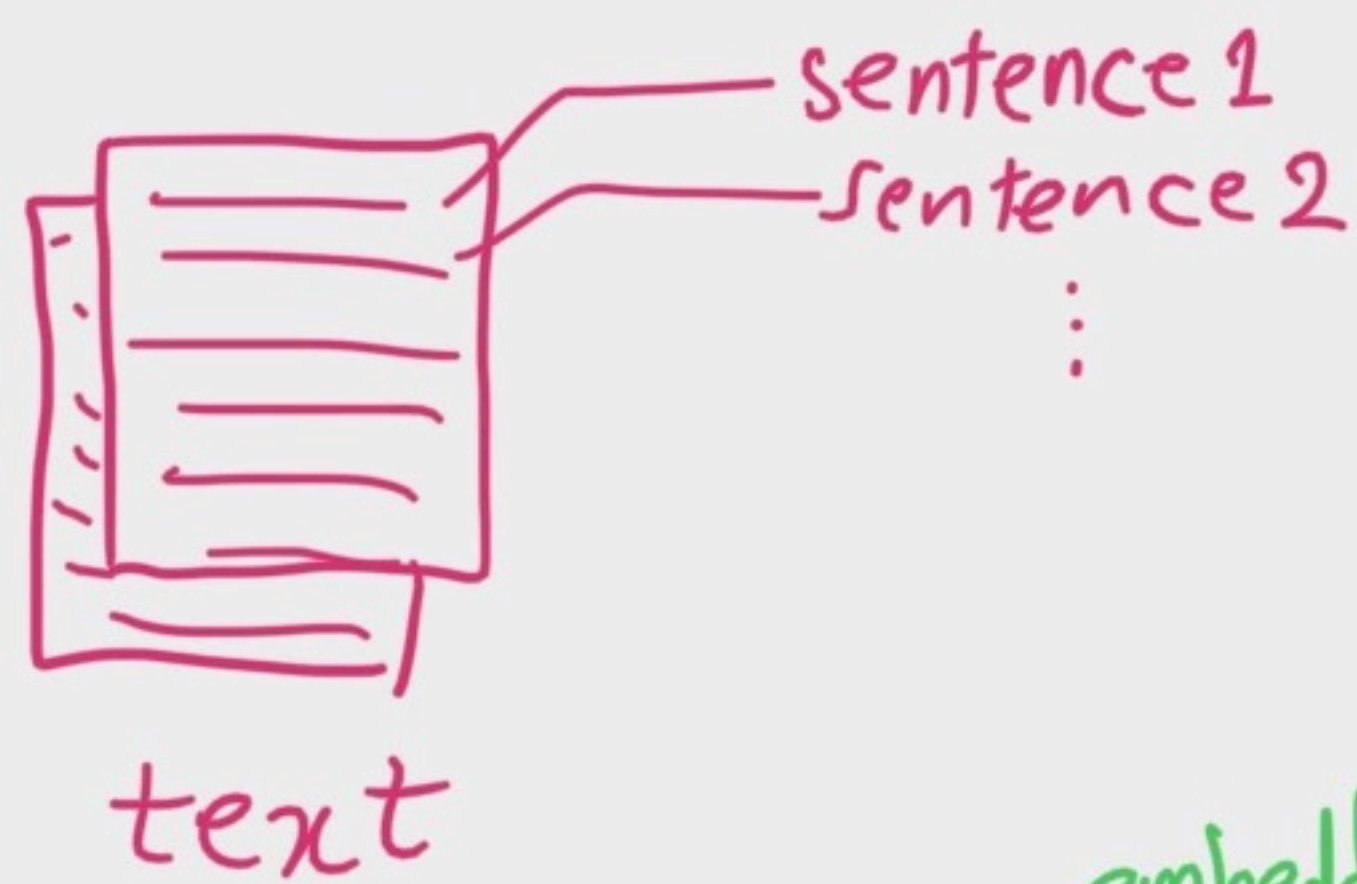
⟶ one-hot encoding is bad!
because dict is huge;
most of cells are zero.

embedding
weight
matrix

input   embedding   hidden   output
layer   layer
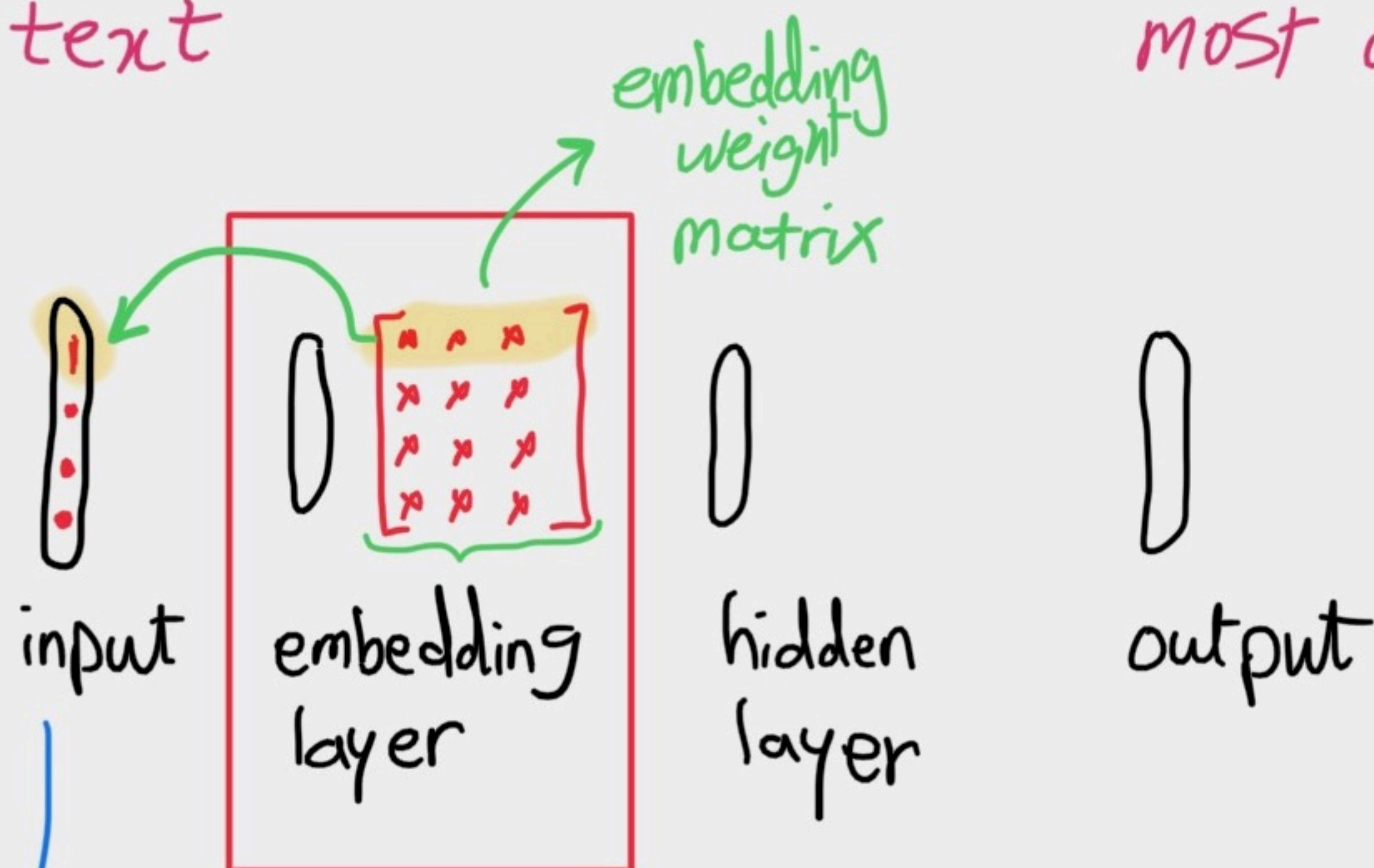
it's a FC layer
with learned embedding weights

we can use a lookup table instead
of this one-hot vector. e.g., represent this with
its corresponding row in the weight embedding matrix.

⟶ embedding lookup