CapsNet Reading Notes

# Capsules vs. Traditional Neurons

|  |  |  |  |
| --- | --- | --- | --- |
| Input from a low-level  capsule / neuron | | (vector) | (scalar) |
| Operations | Affine Transform (仿射变换) |  | None |
| Weighting |  |  |
| Non-linearity |  |  |
| Output | | Vector | Scalar |

# 4 Steps Happening Inside a Capsule (Intuition)

1. Matrix multiplication (affine transformation) of input vector.
2. Scalar weighting of input vectors.
3. Sum of weighted input vectors.
4. “Squashing” non-linearity.



## Affine Transformation

**Weight matrix**

It encodes important spatial (and perhaps other) relationships between lower level features {eyes, mouth, nose} and higher level features {face}.



**e.g.**  encodes where the face should be according to the detected position of eyes.

## Scalar Weighting

**Essence**: decide how much of the lower-level output is to send to each of the upper-level capsules. The lower level capsule measures which upper-level capsule better accommodates its result, say Capsule . Then it will automatically adjust its weight so that  increases and other weight coefficients decreases. This mechanism is called Dynamic Routing.

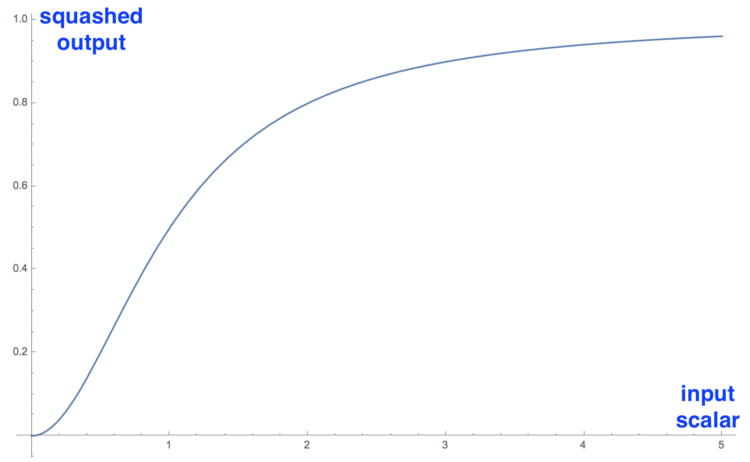
## Summation

Too easy. Omit.

## Squashing

Make the output vector:

* Length: probability (belief) of the existence of a feature.
* Direction: the internal state (e.g. the pose) of the feature.



## Summary

