Regularizing your neural network

Other regularization methods

deeplearning.ai

Data augmentation

Andrew Ng
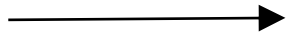
Early stopping

Orthogonalization.

$\rightarrow$ — Optimize cost function J
  — Gradient, ...

$\rightarrow$ — Not overfit.
  — Regularization, ...

L_2

$J(w,b)$

dev set error

training error or J

$w \approx 0$

\# iterations $\|w\|_F^2$

large w