Preliminary Work

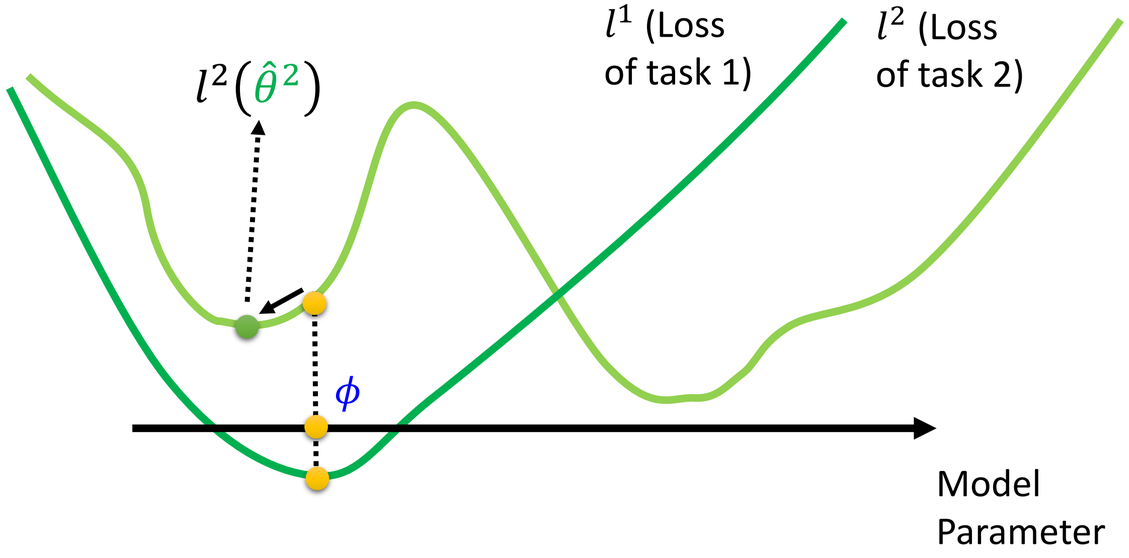
MAML

Model Pre-training

$$L(\phi) = \sum_{n=1}^{N} l^n(\phi)$$

找尋在所有 task 都最好的 ϕ

並不保證拿 ϕ 去訓練以後會得到好的 $\hat{\theta}^n$



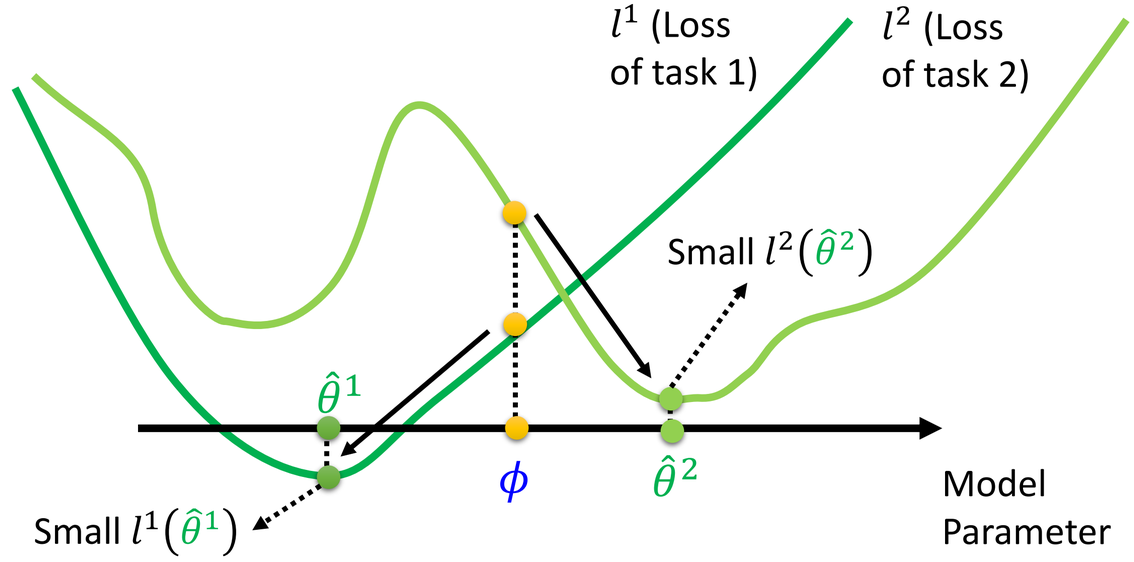
https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Meta1%20(v6).pdf

MAML

$$L(\phi) = \sum_{n=1}^{N} l^n(\hat{\theta}^n)$$

我們不在意 ϕ 在 training task 上表現如何

我們在意用 ϕ 訓練出來的 $\hat{\theta}^n$ 表現如何



https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Meta1%20(v6).pdf

Preliminary Work

MAML

MAML

$\hat{\theta}^n$: model learned from task n

$\hat{\theta}^n$ depends on ϕ

Loss Function:

$$L(\phi) = \sum_{n=1}^{N} l^n(\hat{\theta}^n)$$

$l^n(\hat{\theta}^n)$: loss of task n on the testing set of task n

How to minimize $L(\phi)$? Gradient Descent

$$\phi \leftarrow \phi - \eta \nabla_\phi L(\phi)$$

Find ϕ achieving good performance **after training** 潛力

Model Pre-training

Widely used in transfer learning

Loss Function:

$$L(\phi) = \sum_{n=1}^{N} l^n(\phi)$$

Find ϕ achieving good performance 現在表現如何

https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Meta1%20(v6).pdf