

Model Agnostic Meta Learning (MAML)

Meta RL via gradients

MAML outline

Require $p(\mathcal{T})$: distribution over tasks

Require α, β : step size hyper-parameters

1. randomly initialise θ
2. while not done do
3. sample batch of tasks $\mathcal{T}_i \sim p(\mathcal{T})$
4. for each \mathcal{T}_i do
5. Sample $\mathcal{D}_{\mathcal{T}_i}^{tr} \sim \mathcal{D}_{\mathcal{T}_i}$
6. Sample $\mathcal{D}_{\mathcal{T}_i}^{test} \sim \mathcal{D}_{\mathcal{T}_i}$
7. Evaluate $\nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}_{\mathcal{T}_i}^{tr})$ with respect to K examples
8. Compute adapted parameters with gradient descent: $\phi_i = \theta - \alpha \nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}_{\mathcal{T}_i}^{tr})$
9. Update $\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_i \mathcal{L}(\phi_i, \mathcal{D}_{\mathcal{T}_i}^{test})$

