Language Model(774M) loss curve GPT-method2 3.4 GPT-method5-scale GPT-sincos-method4.2-scale LLaMA-method2 3.2 Validation Loss 2.8 2.6

40

60

Number of iterations(k)

80

100

0

20