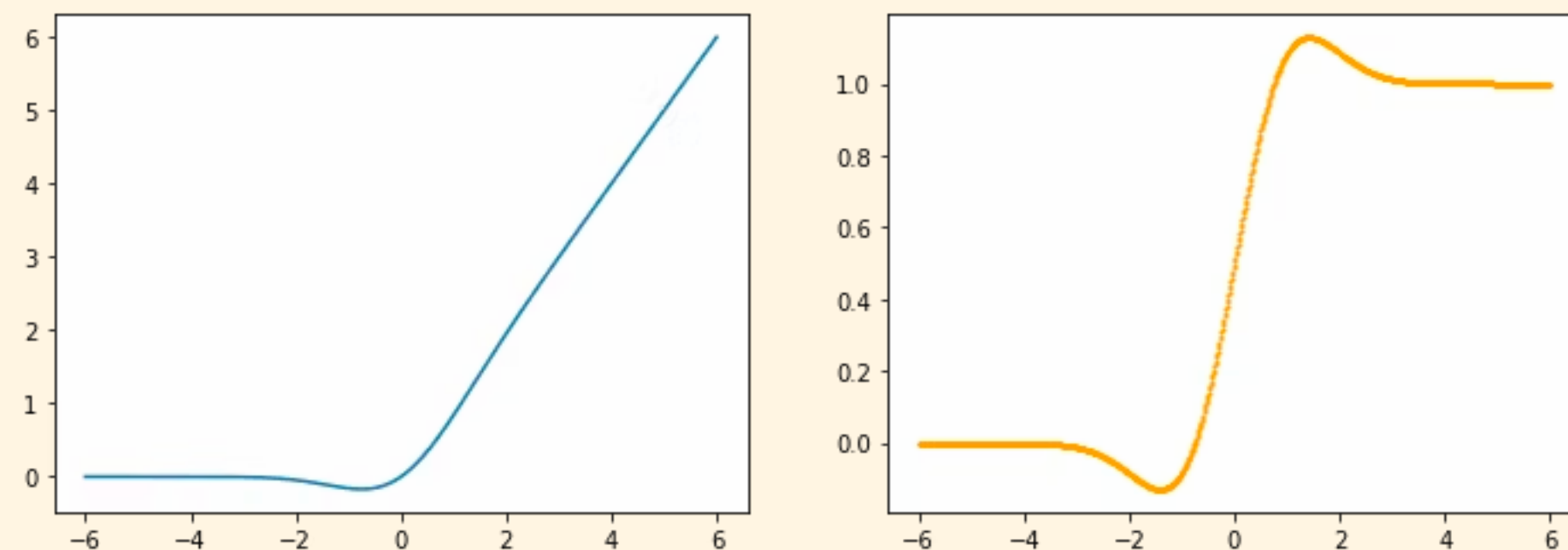
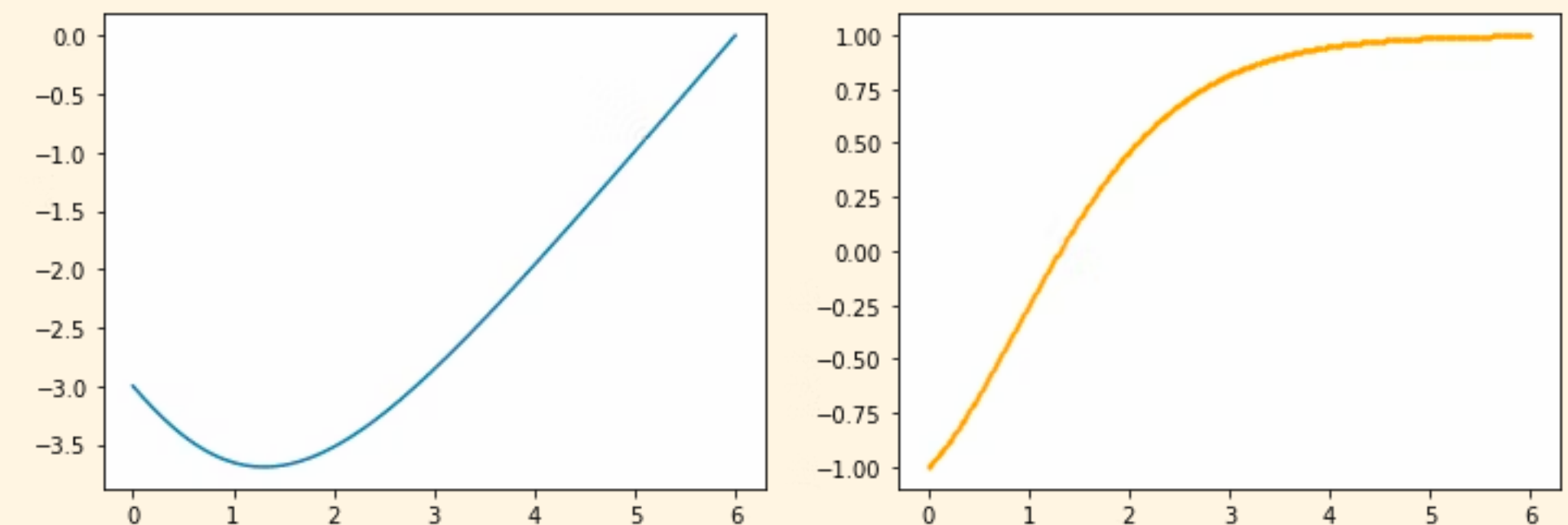


Transformerの構成要素：活性化関数

- 活性化関数は、ニューラルネットワークの表現力を高めるために一般的に用いられる
- 既存のLLMでは、GeLU活性化関数が広く使われている
- 最新のLLM (PaLMやLaMDA) では、GLU活性化関数の系統も使用されており、特にSwiGLUとGeGLUの変種は良い性能を示す



GeLU



GLU

活性化関数とは

