优先拟合低频分量

> [Training behavior of deep neural network in frequency domain]
>
> [Understanding training and generalization in deep learning by Fourier analysis]

任意实空间函数 $f(x)$ 傅里叶得倒空间 $\mathcal{F}[f](\gamma)$

真实函数为 $f(x)$，神经网络输出为 $\tilde{f}(x)$

实空间损失函数 $L=\frac{1}{N}\sum_x|f(x)-\tilde{f}(x)|^2w(x)$

倒空间损失函数 $L=\frac{1}{N}\sum_\gamma|\mathcal{F}[f](\gamma)-\mathcal{F}[\tilde{f}](\gamma)|^2w(\gamma)$

Frequency Principle (F-Principle)：先快速拟合低频率（保持高频率较小），后缓慢拟合高频率

F-Principle can be used to understand the following important phenomena:

- the behavior of DNN training in the information plane, specifically, we focus on understanding the evolution of the mutual information between DNN inputs and DNN outputs during training
- why DNNs capable of overfitting often generalize well
- early-stopping can help avoid fitting the noisy high-frequency components

低频训练集（有意义），高频率较少即幅值较低 高频训练集（噪音），高频率易污染

We can predict that, in the case of insufficient training data, when the higher-frequency components are not negligible, e.g., there exists a significant frequency peak above the effective frequency range, the DNN cannot generalize well after training.