

# Understand FID: Step 1

## 1. 核心问题：如何评价生成图像的质量？

假设你训练了一个模型生成人脸，现在需要判断生成的人脸是否逼真。最直接的方法是：

- **方法A**：找人肉眼观察，但主观性强且效率低。
- **方法B**：用数学公式量化“真实图像”和“生成图像”的差异。

FID 属于方法B，但如何设计这个数学公式？

## 2. 关键直觉：比较“特征分布”而非像素

人类的视觉系统不会逐像素对比两张图，而是提取高级特征（如五官、纹理）。FID模拟了这一过程：

1. **特征提取**：用Inception-v3模型（预训练好的CNN）将图像转换为高级特征（例如2048维向量）。
  - 真实图像 → 特征分布A
  - 生成图像 → 特征分布B
2. **分布比较**：计算两个分布的距离。

**为什么不用像素直接比较？**

像素级对比会忽略语义信息（例如两张人脸五官相同但位置偏移几个像素，人类认为相似，但像素级误差可能很大）。

**FID(A,B)和FID(B,A)相等吗？**

相等。

## 3. FID 具体怎么算？

假设我们有：

- **2张真实图片**（真实分布），经过Inception-v3提取特征后得到：
  - 图片1的特征向量：[1.0, 2.0, 3.0]（假设特征维度=3，简化计算）
  - 图片2的特征向量：[2.0, 3.0, 1.0]
- **2张生成图片**（生成分布）：
  - 生成图片1的特征向量：[1.5, 1.5, 2.5]
  - 生成图片2的特征向量：[2.5, 2.5, 1.5]

**特征矩阵：**

- 真实分布  $\mathbf{X}_r = \begin{bmatrix} 1.0 & 2.0 & 3.0 \\ 2.0 & 3.0 & 1.0 \end{bmatrix}$  (2×3矩阵)
- 生成分布  $\mathbf{X}_g = \begin{bmatrix} 1.5 & 1.5 & 2.5 \\ 2.5 & 2.5 & 1.5 \end{bmatrix}$

首先计算均值。均值是**所有样本在每个特征维度上的平均值**，结果是一个**向量**。

## 计算真实分布的均值 $\mu_r$

$$\mu_r = \frac{1}{2} [1.0 + 2.0 \quad 2.0 + 3.0 \quad 3.0 + 1.0] = [1.5 \quad 2.5 \quad 2.0]$$

## 生成分布的均值 $\mu_g$

$$\mu_g = \frac{1}{2} [1.5 + 2.5 \quad 1.5 + 2.5 \quad 2.5 + 1.5] = [2.0 \quad 2.0 \quad 2.0]$$

然后计算协方差。协方差描述不同特征维度之间的相关性，结果是一个对称方阵。

## 计算真实分布的协方差 $\Sigma_r$

1. 中心化数据（减去均值）：

$$\mathbf{X}_r - \mu_r = \begin{bmatrix} 1.0 - 1.5 & 2.0 - 2.5 & 3.0 - 2.0 \\ 2.0 - 1.5 & 3.0 - 2.5 & 1.0 - 2.0 \end{bmatrix} = \begin{bmatrix} -0.5 & -0.5 & 1.0 \\ 0.5 & 0.5 & -1.0 \end{bmatrix}$$

2. 计算协方差：

$$\begin{aligned} \Sigma_r &= \frac{1}{n-1} (\mathbf{X}_r - \mu_r)^T (\mathbf{X}_r - \mu_r) = \frac{1}{1} \begin{bmatrix} -0.5 & 0.5 \\ -0.5 & 0.5 \\ 1.0 & -1.0 \end{bmatrix} \begin{bmatrix} -0.5 & -0.5 & 1.0 \\ 0.5 & 0.5 & -1.0 \end{bmatrix} \\ &= \begin{bmatrix} (-0.5)^2 + 0.5^2 & (-0.5)(-0.5) + 0.5 \times 0.5 & (-0.5)(1.0) + 0.5 \times (-1.0) \\ \text{同理} & \dots & \dots \\ \text{同理} & \dots & \dots \end{bmatrix} \\ &= \begin{bmatrix} 0.5 & 0.5 & -1.0 \\ 0.5 & 0.5 & -1.0 \\ -1.0 & -1.0 & 2.0 \end{bmatrix} \end{aligned}$$

## 生成分布的协方差 $\Sigma_g$

同理可得：

$$\Sigma_g = \begin{bmatrix} 0.5 & 0.5 & -0.5 \\ 0.5 & 0.5 & -0.5 \\ -0.5 & -0.5 & 0.5 \end{bmatrix}$$

## 计算两个分布的 FID 距离

$$\text{FID} = \|\mu_r - \mu_g\|^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g})$$

1. 中心距离： $\|\mu_r - \mu_g\|^2$ （两分布中心多远）。

2. 形状差异： $\text{Tr}(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g})$ （分布形状多相似）。