

# Group Equivariant Convolutional Networks

## 机器学习课程汇报

何瑞杰

2025-10-29

中山大学 · 数学学院

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

- Title: Group Equivariant Convolutional Networks
- Authors: Taco S. Cohen and Max Welling
- Conference: ICML 2016
- Link: <https://arxiv.org/abs/1602.07576>

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

## 1. 何为等变，等变何为

考虑线性空间  $V$  和上面的一个变换群  $\mathfrak{G}$ ，我们称之为  $\mathfrak{G}$ -空间。对于线性空间上的一个函数  $\Phi$ ，如果它满足

$$\Phi(T_g x) = T'_g \Phi(x)$$

其中  $T_g$  是指对  $V$  中的向量做对应于群元  $g$  的变换。 $T$  和  $T'$  不必相同，但必须要是  $\mathfrak{G}$  中元素的线性表示，即满足对任意  $g, h \in \mathfrak{G}$ ，有  $T(gh) = T(g)T(h)$ 。

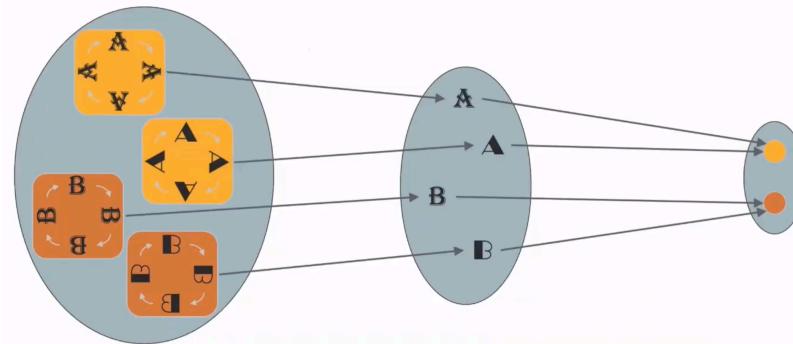


Figure 1: 人眼的旋转等变性——对等变模型的期望

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

## 2. 对称群

典型的对称群分别为  $p4$  和  $p4m$ 。前者是包含了  $\mathbb{Z}^2$  上的所有平移变换和以  $\frac{\pi}{2}$  为单位的旋转变换，它有下面的表示：

$$\mathfrak{g}(r, u, v) = \begin{bmatrix} \cos\left(\frac{r\pi}{2}\right) & -\sin\left(\frac{r\pi}{2}\right) & u \\ \sin\left(\frac{r\pi}{2}\right) & \cos\left(\frac{r\pi}{2}\right) & v \\ 0 & 0 & 1 \end{bmatrix}$$

其中  $r \in \{0, 1, 2, 3\}$ ,  $(u, v) \in \mathbb{Z}^2$ 。群元作用在向量上的结果就可以写为

$$\mathfrak{g}x \simeq \begin{bmatrix} \cos\left(\frac{r\pi}{2}\right) & -\sin\left(\frac{r\pi}{2}\right) & u \\ \sin\left(\frac{r\pi}{2}\right) & \cos\left(\frac{r\pi}{2}\right) & v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}$$

$p4m$  也有类似的表示。

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

### 3. 特征图的信号视角

对于形状为  $[c, w, h]$  的特征图  $F$ , 为推到方便起见, 我们可以将其看作是有一个有界支撑集的函数:

$$F : \mathbb{Z}^2 \rightarrow \mathbb{R}^c$$

$$(x, y) \mapsto \begin{cases} F[:, x, y], & x \in [0, w - 1], y \in [0, h - 1] \\ 0, & \text{otherwise} \end{cases}$$

这样就可以建立和  $\mathbb{R}$  上函数之间卷积操作类似的表达式。对于一个信号  $F$ , 我们定义群元素  $g$  作用在其上的结果为

$$L_g f(x) = f(g^{-1}x).$$

其中  $L_g$  是对应于群元  $g$  之变换  $T_g$  的一个实例化, 并满足  $L_g L_h = L_{gh}$ 。它的直观理解是, 假如  $L_g$  是一个向左的平移变换, 则  $L_g f(x)$  就将信号(例如图片)向左平移  $c$ , 则平移后图片中给定位置的像素值就等于原图片中向右平移相同距离的像素值, 也即  $f(x + c)$ 。

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

## 4.1 通常卷积模块的等变性

首先回忆 CNN 中的卷积操作  $*$  和相关操作  $\star$ ，它们在 CNN 的前向传播和反向传播中成对出现。考虑特征图  $f : \mathbb{Z}^2 \rightarrow \mathbb{R}^{K^{(l)}}$  和一组中的某个卷积核  $\psi^{(i)} : \mathbb{Z}^2 \rightarrow \mathbb{R}^{K^{(l)}}$ ，有

$$[f * \psi](\mathbf{x}) = \sum_{\mathbf{y} \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(\mathbf{y}) \psi_k^{(i)}(\mathbf{x} - \mathbf{y}) \quad \text{convolution}$$

$$[f \star \psi](\mathbf{x}) = \sum_{\mathbf{y} \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(\mathbf{y}) \psi_k^{(i)}(\mathbf{y} - \mathbf{x}) \quad \text{correlation}$$

## 4.1 通常卷积模块的等变性

现在我们验证相关操作的平移等变性。考虑  $\mathbb{Z}^2$  上的平移群元  $t$  对应的变换  $L_t$ , 有

$$\begin{aligned}
 [(L_t f) \star \psi](x) &= \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} [L_t f_k](y) \psi_k^{(i)}(y - x) \\
 &= \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(y - t) \psi_k^{(i)}(y - x) \\
 &= \sum_{z \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(z) \psi_k^{(i)}(z - (x - t)) \\
 &= [f \star \psi](x - t) = [L_t[f \star \psi]](x). \quad z \leftarrow y - t
 \end{aligned}$$

## 4.1 通常卷积模块的等变性

相关变换对旋转没有等变性，对于  $p4$  中的群元  $r$  对应的旋转变换  $L_r$ ，有：

$$\begin{aligned}
 [(L_r f) \star \psi](x) &= \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} [L_r f_k](y) \psi_k^{(i)}(y - x) \\
 &= \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(r^{-1}y) \psi_k^{(i)}(y - x) \\
 &= \sum_{z \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(z) \psi_k^{(i)}(rz - x) \\
 &= \sum_{z \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(z) \psi_k^{(i)}(r(z - r^{-1}x)) \quad z \leftarrow r^{-1}y
 \end{aligned}$$

$$\begin{aligned} &= \sum_{z \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(z) \left[ L_{r^{-1}} \psi_k^{(i)} \right] (z - r^{-1}x) \\ &= [f \star L_{r^{-1}} \psi](r^{-1}x) = L_r[f \star L_{r^{-1}} \psi](x) \end{aligned}$$

可见通常的相关操作对旋转没有等变性。对于卷积，注意只需令  $\varphi(x) = \psi(-x)$ ，相关操作就变成了卷积操作。因此卷积的等变性和相关操作相同。

## 4.2 群等变相关操作

为了让卷积操作对旋转、以至更加一般的操作具有等变性，作者提出了群相关操作。注意上文中的相关操作

$$[f \star \psi^{(i)}](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k(y) \psi_k^{(i)}(-x + y)$$

注意  $x$  对应着  $\mathbb{Z}^2$  上所有平移操作构成的群中的一个群元素  $g$ ，而  $-x$  对应着它的逆元  $g^{-1}$ 。

## 4.2 群等变相关操作

我们可以自然地将原来相关操作写成包含群元素的形式，这就得到了第一层的群相关：

$$f^{(1)}(\mathbf{g}) = [f \star \psi^{(i)}](\mathbf{g}) = \sum_{\mathbf{y} \in \mathbb{Z}^2} \sum_{k=1}^{K^{(l)}} f_k^{(0)}(\mathbf{y}) \psi_k^{(0,i)}(\mathbf{g}^{-1} \mathbf{y})$$

其中  $\mathbf{g} \in \mathfrak{G}$ ,  $f$  和  $\psi^{(0,i)}$  都是  $\mathbb{Z}^2$  到  $\mathbb{R}$  的函数，但得到的新信号  $[f \star \psi^{(i)}]$  的定义域为  $\mathfrak{G}$ 。因此接下来的各层中群相关操作的表达式需要做一些微调：

$$f^{(l+1)}(\mathbf{g}) = [f \star \psi^{(i)}](\mathbf{g}) = \sum_{\mathbf{h} \in \mathfrak{G}} \sum_{k=1}^{K^{(l)}} f_k^{(l)}(\mathbf{h}) \psi_k^{(l,i)}(\mathbf{g}^{-1} \mathbf{h})$$

其中  $l \geq 1$ ,  $\mathbf{g}, \mathbf{h} \in \mathfrak{G}$ ,  $f_k^{(l)}$  和  $\psi_k^{(l,i)}$  的定义域都是  $\mathfrak{G}$  (这里假设每层的变换群相同)。

## 4.2 群等变相关操作

接下来证明它是关于群  $\mathfrak{G}$  的元素等变的:

$$\begin{aligned}
 [[L_{\mathbf{u}} f] \star \psi](\mathbf{g}) &= \sum_{\mathbf{h} \in X} \sum_{k=1}^K [L_{\mathbf{u}} f_k](\mathbf{h}) \psi_k^{(i)}(\mathbf{g}^{-1} \mathbf{h}) \\
 &= \sum_{\mathbf{h} \in X} \sum_{k=1}^K f_k(\mathbf{u}^{-1} \mathbf{h}) \psi_k^{(i)}(\mathbf{g}^{-1} \mathbf{h}) \\
 &= \sum_{\mathbf{p} \in X} \sum_{k=1}^K f_k(\mathbf{p}) \psi_k^{(i)}(\mathbf{g}^{-1} \mathbf{u} \mathbf{p}) \\
 &= \sum_{\mathbf{p} \in X} \sum_{k=1}^K f_k(\mathbf{p}) \psi_k^{(i)}\left((\mathbf{u}^{-1} \mathbf{g})^{-1} \mathbf{p}\right) \quad \mathbf{p} \leftarrow \mathbf{u}^{-1} \mathbf{h} \\
 &= [f \star \psi](\mathbf{u}^{-1} \mathbf{g}) = [L_{\mathbf{u}}[f \star \psi]](g)
 \end{aligned}$$

## 4.2 群等变相关操作

注意当  $\mathfrak{G}$  不是交换群时，群卷积和群相关操作也不交换，但是有  $f \star \psi = [\psi \star f]^*$ ，其中  $\square^*$  是内卷积操作，即  $f^*(g) = f(g^{-1})$ ：

$$\begin{aligned}
 [f \star \psi](\mathfrak{g}) &= \sum_{\mathfrak{h} \in X} \sum_{k=1}^K f_k(\mathfrak{h}) \psi_k(\mathfrak{g}^{-1} \mathfrak{h}) \\
 &= \sum_{\mathfrak{h} \in X} \sum_{k=1}^K f_k(\mathfrak{h}) \psi_k(\mathfrak{g}^{-1} \mathfrak{h}) \quad \mathfrak{p} \leftarrow \mathfrak{g}^{-1} \mathfrak{h} \\
 &= \sum_{\mathfrak{p} \in X} \sum_{k=1}^K \psi_k(\mathfrak{p}) f_k\left((\mathfrak{g}^{-1})^{-1} \mathfrak{p}\right) \\
 &= [\psi \star f](\mathfrak{g}^{-1}) = [\psi \star f]^*(\mathfrak{g}).
 \end{aligned}$$

## 4.3 线性组合和单点函数的等变性

假设有两个关于群  $\mathfrak{G}$  等变的信号  $f(\cdot)$  和  $g(\cdot)$ , 显然其线性组合也是关于群  $\mathfrak{G}$  等变的:

$$[L_{\mathfrak{g}}[af + bg]](\mathfrak{u}) = [af + bg](\mathfrak{g}^{-1}\mathfrak{u}) = [aL_{\mathfrak{g}}f + bL_{\mathfrak{g}}g](\mathfrak{u})$$

其中  $\mathfrak{g} \in \mathfrak{G}$ 。对于单点非线性函数  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ , 其作用在信号上的方式是简单的函数复合, 即  $\sigma \circ f$  容易证明它对群  $\mathfrak{G}$  中的任一元素  $\mathfrak{g}$  的等变性:

$$L_{\mathfrak{g}}[\sigma \circ f](\mathfrak{u}) = [\sigma \circ f](\mathfrak{g}^{-1}\mathfrak{u}) = \sigma[f(\mathfrak{g}^{-1}\mathfrak{u})] = [\sigma \circ L_{\mathfrak{g}}](\mathfrak{u}).$$

## 4.4 子群池化和陪集池化

回忆通常卷积网络的（最大值）池化操作，对于单通道的信号  $f: \mathbb{Z}^2 \rightarrow \mathbb{R}$ ，它其实可以写为

$$[\text{MaxPool}_1 f](x) = \max_{u \in \mathcal{N}(x)} f(u)$$

其中  $\mathcal{N}(x)$  是  $x$  的一个邻域，实践中常常为以  $x$  为中心的一个正方形邻域，如九宫格。再在  $\mathbb{Z}^2$  的某个子集（例如  $(3\mathbb{Z})^2$ ）上采样，最后再映射回  $\mathbb{Z}^2$ 。这就是池化中的两步操作：

1. 计算给定位置  $x$  的邻域最大值
2. 选取一部分邻域最大值作为输出

## 4.4 子群池化和陪集池化

我们可以依据这两步构造包含群的版本。首先对于第一步, 我们可以考虑一个在单位变换  $\mathbf{e}$  的一个包含于群  $\mathfrak{G}$  的“邻域”  $\mathfrak{U} \in \mathfrak{G}$ , 然后考虑  $\mathbf{g}\mathfrak{U} := \{\mathbf{g}\mathbf{u} : \mathbf{u} \in \mathfrak{U}\}$ , 这就得到了邻域最大值的群元素版本:

$$Pf(\mathbf{g}) = \max_{\mathbf{k} \in \mathbf{g}\mathfrak{U}} f(\mathbf{k})$$

这个操作是群等变的:

$$\begin{aligned} [L_{\mathbf{u}} P]f(\mathbf{g}) &= L_{\mathbf{u}}[Pf](\mathbf{g}) = [Pf](\mathbf{u}^{-1}\mathbf{g}) = \max_{\mathbf{k} \in \mathbf{u}^{-1}\mathbf{g}\mathfrak{U}} f(\mathbf{k}) \\ &= \max_{\mathbf{h} \in \mathbf{g}\mathfrak{U}} f(\mathbf{u}^{-1}\mathbf{h}) = \max_{\mathbf{h} \in \mathbf{g}\mathfrak{U}} [L_{\mathbf{u}}f](\mathbf{h}) \quad \mathbf{h} \leftarrow \mathbf{u}\mathbf{k} \\ &= [PL_{\mathbf{u}}](\mathbf{g}). \end{aligned}$$

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

## 5. 群相关操作的高效实现

对于平面的变换群  $\mathfrak{G}$ , 如果每个群元  $g \in \mathfrak{G}$  都可以写成一个平移变换  $t \in \mathbb{Z}^2$  和一个稳定子, 即保零点的旋转变换  $s$  的复合, 我们称它是可分的。例如群  $p4$ , 它的元素  $g$  总是可以写成  $g = ts$ , 其中  $t$  是平移,  $s$  是以原点为旋转中心的旋转, 根据前文中提到的群元对应的变换的同态性质, 有

$$[f \star \psi^{(i)}](g) = [f \star \psi^{(i)}](ts) = \sum_{h \in \mathfrak{X}} \sum_k f_k(h) L_t \left[ L_s \psi_k^{(i)}(h) \right]$$

注意平移变换  $L_t$  对应的就是通常卷积中的平移变换。所以可以先算出  $L_s \psi_k$ , 然后直接套用通常的卷积即可。其次, 注意这里的  $\psi_k$  是群  $\mathfrak{G}$  到  $\mathbb{R}$  的映射, 因此该映射可以直接使用一个数组  $F$  来表示——我们不需要群元素的表示, 而是只需要将它们编号, 然后构造一个字典/数组来查找即可。

## 5. 群相关操作的高效实现

考虑  $\mathbb{Z}^2$  上可分的变换群  $\mathfrak{G}$ ，那么该群一共有  $S \times n \times n$  个群元。假设上一层的通道数为  $K$ ，下一层通道数为  $K'$ ，那么对应于这一层卷积核的数组  $\mathbf{F}$  的形状为

$$\overbrace{K' \times K \times \underbrace{S \times n \times n}_{\substack{\text{群元素编号, 对应于 } \psi_k^{(i)} \\ \text{考虑上一层的所有通道, 对应于 } \psi^{(i)}}}}^{\substack{\text{考虑下一层的所有通道, 对应于完整的 } \psi}}.$$

接着我们需要将  $\psi$  和  $L_{\mathfrak{s}}$  配对，由于  $\mathfrak{s}$  是任选的稳定子，配对后相当于一个新的函数  $g(\mathfrak{s}', \mathfrak{s}, \mathfrak{t}) = h(\mathfrak{s}'^{-1}\mathfrak{s}\mathfrak{t})$ ，所以配对后的结果是  $K' \times S \times K \times S \times n \times n$ ，我们将配对后的数组记为  $\mathbf{G}$ 。

## 5. 群相关操作的高效实现

为了计算  $s'^{-1}st$  我们需要用到一个双射  $g(s, u, v)$ , 它接受  $S \times n \times n$  中的一个三元组, 这对应于群  $\mathfrak{G}$  中一个元素的指标, 输出该群元素对应的(可逆的)矩阵表示。因此  $st$  对应的指标是  $(s, t_x, t_y)$ , 而  $s'$  由于是稳定子, 没有复合平移变换, 因此对应的指标是  $(s', 0, 0)$ , 因此可以得到  $s'^{-1}st$  对应的指标是

$$(\bar{s}, \bar{t}_x, \bar{t}_y) = g^{-1}[(g(s', 0, 0))^{-1}g(s, t_x, t_y)]$$

因此就有对应  $G[i, s', j, s, t_x, t_y] = F[i, j, \bar{s}, \bar{t}_x, \bar{t}_y]$ 。这样得到的  $G$  就是对应于  $L_s\psi$  的映射字典。

接下来要执行的是通常的卷积操作, 只需将  $G$  的形状  $K' \times S' \times K \times S \times n \times n$  变形为  $(K'S) \times (KS) \times n \times n$ , 即输入通道数为  $KS$  输出通道数为  $K'S$  的卷积; 特征图  $f$  也可以做类似的操作, 其原本的形状为  $K \times S \times n \times n$ , 将前两维合并, 得到的形状是  $(KS) \times n \times n$ , 和改变形状后的卷积核对应。

# Outline

1. 何为等变，等变何为 .....	3
2. 对称群 .....	5
3. 特征图的信号视角 .....	7
4. 等变性 .....	9
4.1 通常卷积模块的等变性 .....	10
4.2 群等变相关操作 .....	14
4.3 线性组合和单点函数的等变性 .....	18
4.4 子群池化和陪集池化 .....	19
5. 群相关操作的高效实现 .....	21
6. 实验结果和讨论 .....	25

## 6. 实验结果和讨论

实验中用作参照的模型是全卷积网络和 ResNet，在做了随机旋转的 RotatedMNIST 和 CIFAR 系列数据集上都有相比于其他模型更好的效果；在群卷积的范畴之中（通常卷积可以视作  $\mathbb{Z}^2$  上平移变换群卷积这一特殊情形）拥有最多群元素  $p4m$  在 CIFAR 系列数据集上的效果最好。

*Table 1:* 本文模型 (P4CNN) 与各 *Baseline* 模型在 *Rotated MNIST* 上的训练误差

Network	Test Error (%)
Larochelle et al. (2007)	$10.38 \pm 0.27$
Sohn & Lee (2012)	4.2
Schmidt & Roth (2012)	3.98
Z2CNN	$5.03 \pm 0.0020$
P4CNNT Rotation Pooling	$3.21 \pm 0.0012$
<b>P4CNN</b>	<b><math>2.28 \pm 0.0004</math></b>

*Table 2:* 通常卷积网络 ( $\mathbb{Z}^2$ )、 $p4$  和  $p4m$  在 CIFAR10 和 CIFAR10+ 上的训练误差和模型参数量

Network	$G$	CIFAR10	CIFAR10+	Param.
All-CNN	$\mathbb{Z}^2$	9.44	8.86	1.37M
	$p4$	8.84	7.67	1.37M
	$p4m$	7.59	7.04	1.22M
ResNet44	$\mathbb{Z}^2$	9.45	5.61	2.64M
	$p4m$	<b>6.46</b>	<b>4.94</b>	2.62M

## 6. 实验结果和讨论

从上面的结果可以看到，群等变模块可以替换通常卷积模型中的卷积、非线性层和池化层，并得到更好的效果。

群等变 CNN 的进一步的拓展改进包括将六边形网格中的变换群、 $\mathbb{R}^3$  中的变换群、连续（局部紧）群以及更大的有限群。不过对于后两个实现难度和计算成本较大。