problems :-

lec 2

- Lower bound of generalized error for lipz functions
Not solved !

Lec 3

① The inner product satisfy axioms ?

solution: The axioms are:

$$\langle ax, y\rangle = a\langle x, y\rangle$$

$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$$

$$\langle x, x\rangle > 0$$

$$\langle ax, y\rangle = \int_{\Omega} \langle ax(\omega), y(\omega)\rangle_C \, d\mu(\omega)$$

$$= a\int_{\Omega} \langle x(\omega), y(\omega)\rangle_C \, d\mu(\omega)$$

$$= a\langle x, y\rangle$$

and the remaining axioms can be proved in a similar manner

② can you think of other weight space symmetries ?

change the sign of the weight & bias Matrices for the 1^{st} & 2^{nd} Layer in the MLP, taking an odd activation Function like (tanh)

③

Scanned with CamScanner

③ Explain why if we Knew all symmetries of L, we only need 1 label per class?

because if we Know all symmetries of L that includes all transformations, we can generate all possible examples from only one example.

④ Verify that ~~Euclic~~ Euclidean planar motions ~~def~~ is a group action, defined by

$$((\theta, t_x, t_y), (x,y)) \rightarrow \begin{bmatrix} \cos\theta & \sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

two group ~~action~~ is composed as

$$(\theta_1, t_{x_1}, t_{y_1}) \circ (\theta_2, t_{x_2}, t_{y_2}) = (\theta_1+\theta_2, t_{x_1}+t_{x_2}, t_{y_1}+t_{y_2})$$

to show ~~[illegible]~~ it's a group it should satisfy: $\forall g,h \in G \quad (gh)u = g(hu)$

$\exists e \in G \quad eu = u$

let $g = (\theta_1, t_{x_1}, t_{y_1}) \qquad h = (\theta_2, t_{x_2}, t_{y_2})$

$$\cancel{g(hu) = g \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & t_x \\ \sin\theta_2 & \cos\theta_2 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}$$

$$\cancel{= g \begin{bmatrix} x\cos\theta_2 + y\sin\theta_2 + t_{x_2} \\ x\sin\theta_2 + y\cos\theta_2 + t_{y_2} \\ 1 \end{bmatrix}}$$

$$= \begin{bmatrix} \cos\theta_1 & \sin\theta_1 & t_{x_1} \\ \sin\theta_1 & \cos\theta_1 & t_{y_1} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x\cos\theta_2 + y\sin\theta_2 + t_{x_2} \\ x\sin\theta_2 + y\cos\theta_2 + t_{y_2} \\ 1 \end{bmatrix}$$

or Matrix multiplication is associative

$$(gh)u = \left(\begin{bmatrix} \cos\theta_1 & \sin\theta_1 & t_{x_1} \\ \sin\theta_1 & \cos\theta_1 & t_{y_1} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta_2 & \sin\theta_2 & t_{x_2} \\ \sin\theta_2 & \cos\theta_2 & t_{y_2} \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 & t_{x_1} \\ \sin\theta_1 & \cos\theta_1 & t_{y_1} \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} \cos\theta_2 & \sin\theta_2 & t_{x_2} \\ \sin\theta_2 & \cos\theta_2 & t_{y_2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \right)$$

because Matrix Multiplication is associative

$$= g(hu) \quad \#$$

$$\text{for } e = \begin{bmatrix} \cos 0 & \sin 0 & 0 \\ \sin 0 & \cos 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\theta = 0°$ & $t_x = t_y = 0$

$\Rightarrow e = I$ (Identity matrix)

$$\Rightarrow \quad e \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = u$$

$\Rightarrow eu = u$

then (θ, t_x, t_y) is a group action

⑤ for the Group $(\mathbb{Z}, +)$ action $g = n$ $h = m$

proof that $\rho(n)$ is a group representation

$$\rho(n) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}^{n}$$

we have to prove that $\rho(gh) = \rho(g)\rho(h)$

$gh = n + m$ $\quad \rho(gh) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}^{n+m}$

which can be decomposed as $\rho(n)\,\rho(m)$

#

~~Now derive the reps for C4 group on grid 3x3~~

also prove: $\rho(g)\,\rho^{-1}(g) = I$ (invertable)

we find a Left shift matrix & multiply it by the right shift, we should get I

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = I$$

#

⑥ Derive a reps for the group C_4 on 3x3 grid

$C_4 = (\mathbb{Z}_4, +)$ $\quad g = n$ $\quad h = m$ $\quad g \circ h = (n+m) \% 4$

$\Omega = \mathbb{Z}_3 \times \mathbb{Z}_3$

$g = r$ $\quad r \in \{0, 1, 2, 3\}$

$$\rho(g) f(u) = \underbrace{\begin{bmatrix} \cos r\frac{\pi}{2} & \sin r\frac{\pi}{2} \\ \sin r\frac{\pi}{2} & \cos r\frac{\pi}{2} \end{bmatrix}}_{\rho^{\Omega}(g)} \begin{bmatrix} x \\ y \end{bmatrix}$$

7 for each permutation ~~matrix~~ on the Sets (scaler feature, vector feature and Tensor feature) verify it is a group repr.

$P_0(P_{12})s = 1 \cdot s$

$P_1(P_{12})v = P_{12}v$

$P_2(P_{12})M = P_{12}MP_{12}^T$

Scanned with CamScanner

⑧ Show that if f_i are all equivarient, so their composition.

we need to show

$(g \cdot f_{i+1} \circ f_i)(x) = f_{i+1} \circ f_i (g x)$

$(f_{i+1} \circ f_i) \circ f_{i-1}(g)$

$\Rightarrow f_{i+1} \circ (f_i \circ f_{i-1}(g)) = f_{i+1} \circ (f_i(g) \circ f_i)$

$= (f_{i+1} \circ f_i(g)) \circ f_i = f_{i+1}(g) \circ (f_{i+1} \circ f_i)$

we proved equivariance ✱

(9)

Show that convolution is translation equivariant But not rotation equivariant

Convolution is defined as:-

$$[f * \varphi](x) = \sum_{y \in \mathbb{Z}^2} \sum_{k=1} f(y)\,\varphi(y - x)$$

leave out the sum over channels for clarity

put $y \to y + t$

we have to show that translation followed by convolution is the same as convolution followed by translation As:-

$$[[L_t f] * \varphi](x) = \sum_y f(y-t)\,\varphi(y-x)$$

$$= \sum_y f(y)\,\varphi(y+t-x)$$

$$= \sum_y f(y)\,\varphi(y-(x-t)) = (f * \varphi)(x-t)$$

$$= [L_t[f * \varphi]](x)$$

#

2 for Rotation :-

$$[L_r f] * \psi](x) = \sum_y \sum_k L_r f(y) \psi(y-x)$$

$$= \sum_y \sum_k f(r^{-1}y) \psi(y-x)$$

$$= \sum_y \sum_k f(y) \psi(ry-x)$$

$$= \sum_y \sum_k f(y) \psi(r(y - r^{-1}x))$$

$$= \sum_y \sum_k f(y) L_{r^{-1}} \psi(y - r^{-1}x)$$

$$= \cancel{f(y) * [L_{r^{-1}}[\psi](r^{-1}x)]}$$

$$= f * [L_{r^{-1}} \psi](r^{-1}x)$$

$$= L_r [f * [L_{r^{-1}} \psi]](x)$$

$$\therefore [[L_r f] * \psi](x) \neq L_r [f * \psi](x)$$

Lec 4 :

Not Solved !!

$$\Omega = \{1, \dots, d\}$$

$$G = C_d$$

$$F = \{ f(x) = P_k(x_1, \dots, x_d)' \text{ of degree } k \}$$

$$S_G F = ??$$

$$f(x) = \sum_{|\alpha| \le m} c_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

$m = k$

$n = d$

$$|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$$

$$S_G f = \cancel{(g_1 f(x) + g_2 f(x) + \dots g_d f(x))} \frac{1}{d}$$

$$= \frac{1}{d}\left(f(g_1 x) + \dots f(g_d x)\right)$$

Scanned with CamScanner

② why combining linear equivariant with ~~non~~ linear invariants isn't useful for extracting invariants?

because composing two linear function will give you another linear function so our hypothesis will still ~~the~~ be the invariant function which loses information

Lec 5

① To ensure equivariance of $F(X, A)$ it is sufficient that ϕ is permutation invariants in X_N

ϕ invariant $\Rightarrow$ in X_{N_i}

$$\phi(x_i, PX_{N_i}) = \phi(x_i, X_{N_i}) \quad — ①$$

$$F(PX, PAP^T) = P\begin{bmatrix} \phi(x_i, PX_{N_i}) \\ \vdots \\ \phi(x_i, PX_{N_i}) \end{bmatrix} \quad — ②$$

put ① in ②

$$F(PX, PAP^T) = P\begin{bmatrix} \phi(x_i, X_{N_i}) \\ \vdots \\ \phi(x_i, X_{N_i}) \end{bmatrix} = \cancel{PX} \; PF(X, A)$$

② prove that

Convolutional ⊆ attential ⊆ message-passing

On Campus wire