Mini-course on representation theoretical methods in ML

3. Fourier transforms and the symmetric group

Risi Kondor (Gatsby Unit)

1. G acts on a vector space U via the linear operators

$$T_g \colon U \to U \qquad\qquad g \in G$$

which must satisfy $T_{g_1} T_{g_2} = T_{g_1 g_2}$.

2. Equivalent to a system of matrices $\rho\colon G \to \mathbb{C}^{d\times d}$ satisfying

$$\rho(g_1)\,\rho(g_2) = \rho(g_1 g_2).$$

3. Notion of equivalence and reducibility

$$\rho_1(x) = T^{-1}\rho_2(x)\,T \qquad\qquad \rho(x) = T^{-1}\begin{pmatrix}\rho_1(x) & 0 \\ 0 & \rho_2(x)\end{pmatrix} T$$

leads to complete set of inequivalent irreducible representations $\mathcal{R}$.

4. Any $G-$module reduces in the form

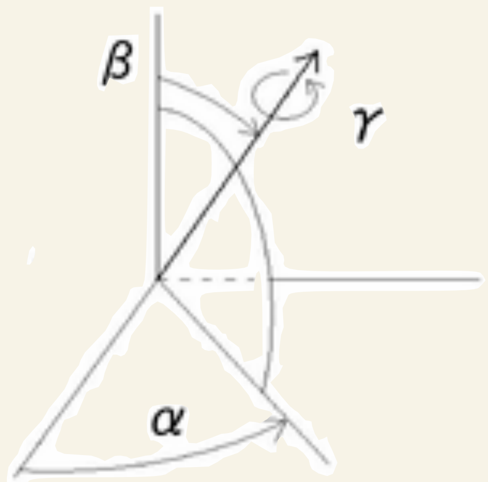
$$U = W_1 \oplus W_2 \oplus \ldots \oplus W_k$$

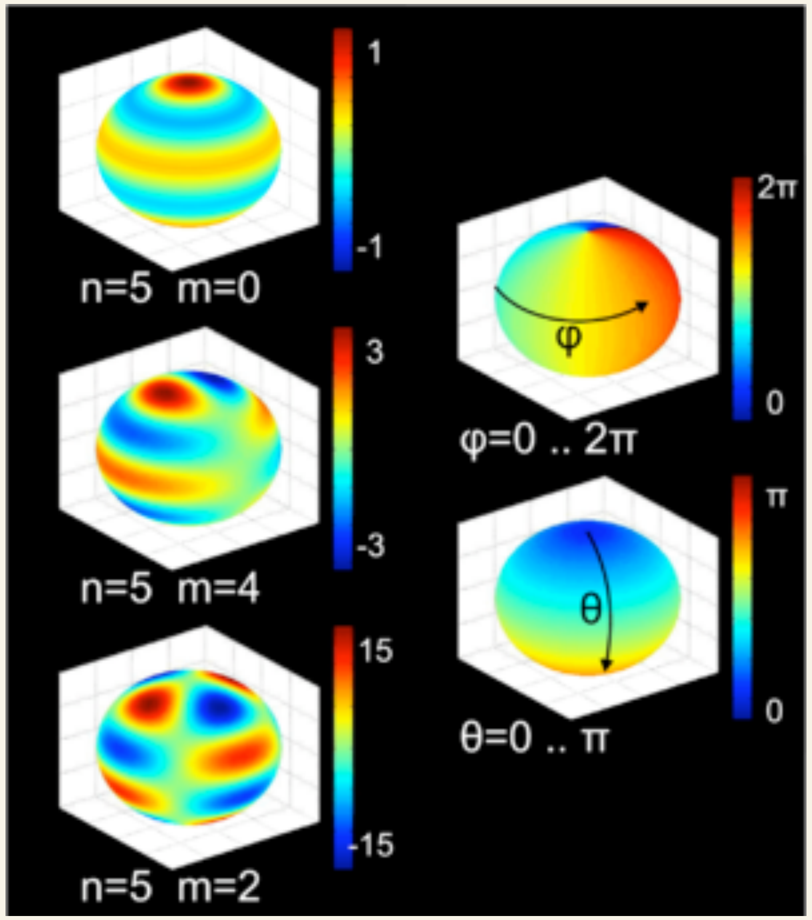
$$T_g(f) = [T_g]_1\,([f]_1) \oplus \ldots \oplus [T_g]_k\,([f]_k)$$

1. $G = \mathbb{R}$ $\rho_k(x) = e^{ikx}$

2. $G = SO(3)$

$$[\rho_l(\theta, \phi, \psi)]_{m,m'} = e^{-il\psi}\, Y_l^m(\theta, \phi) \qquad l = 2k+1$$





5. A specific type of $G-$module:

$$U = L(\mathcal{X}) = \{f \colon \mathcal{X} \to \mathbb{C}\}$$

where G acts on $\mathcal{X}$ by $x \mapsto g(x)$, and by extension

$$f \mapsto f^g \qquad f^g(x) = f(g^{-1}(x)).$$

6. Now what about taking $\mathcal{X} = G$ and $g_1(g_2) = g_1 g_2$?

$$f \mapsto f^y \qquad f^y(x) = f(y^{-1}x).$$

Transformations in Quantum Mechanics:

$$|\psi\rangle \mapsto U(g)|\psi\rangle \qquad\qquad g \in G$$

unitary representation of G

symmetry group $\longrightarrow$ embed in $\mathrm{SU}(n)$ $\longrightarrow$ look at irreps

Example: spin

$\mathrm{SO}(3) \hookrightarrow \mathrm{SU}(2)$ double cover!

$\mathrm{SU}(2)$ has one irrep for each $d \in \mathbb{N}$

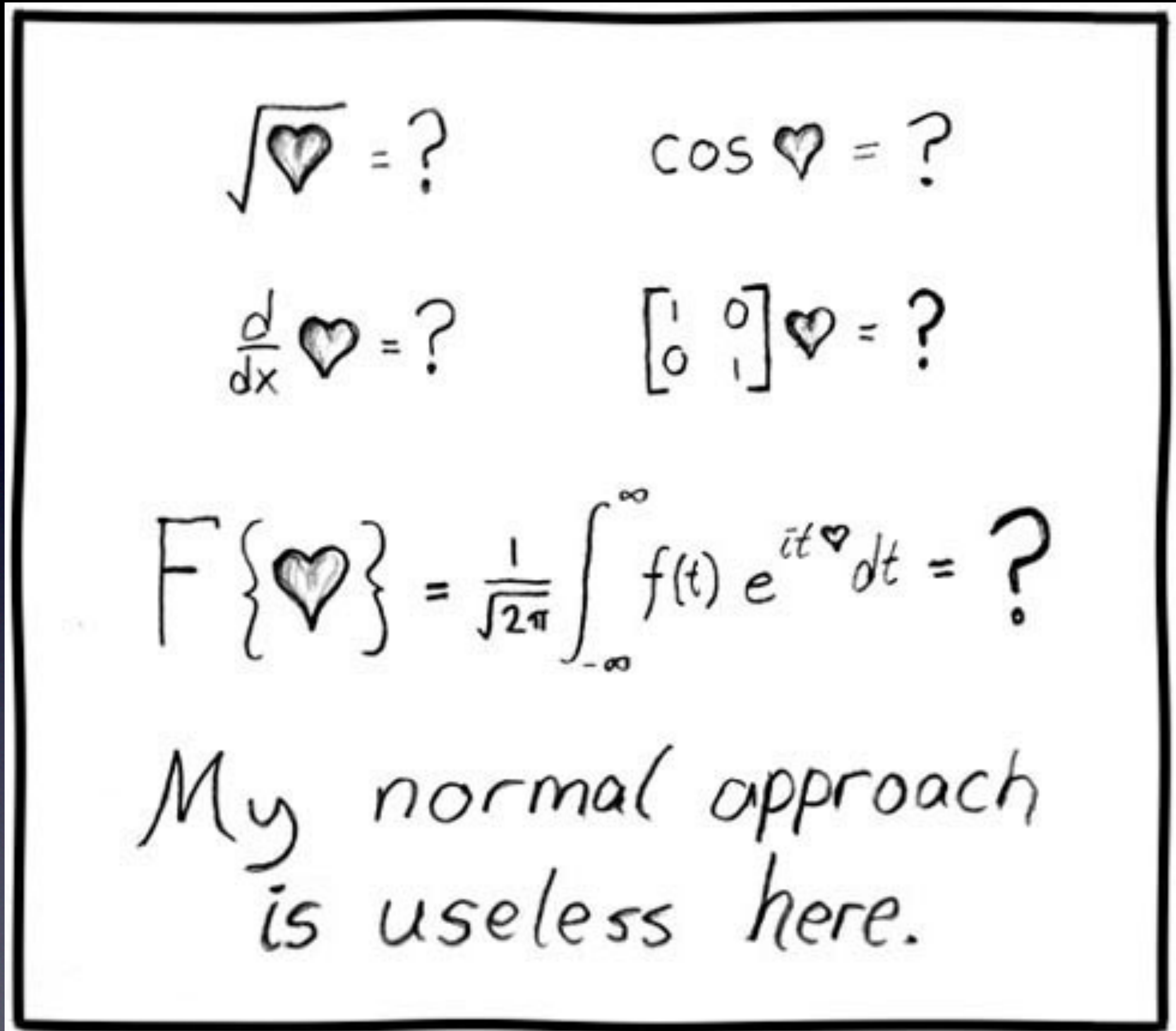
Fermions:

$d = 2s \quad m = -(s-1/2)\hbar, -(s-3/2)\hbar, \ldots, (s-1/2)\hbar$

Bosons:

$d = 2s+1 \qquad m = -s\hbar, -(s-1)\hbar, \ldots, s\hbar$

back to Fourier transforms


√♥ = ?
cos ♥ = ?
d/dx ♥ = ?
[1 0; 0 1] ♥ = ?
F{♥} = 1/√2π ∫ from -∞ to ∞ f(t) e^(it♥) dt = ?
My normal approach is useless here.

xkcd.com

G acts on $L(G)$ by $f \mapsto f^y$, $f^y(x) = f(y^{-1}x)$.

Theorem(Peter-Weyl). The decomposition of $L(G)$ into irreducible G-modules contains each irreducible G-module with multiplicity equal to its degree.

$$\mathrm{L(G)} = \bigoplus_{\rho \in \mathcal{R}} \underbrace{\bigoplus_{i=1}^{d_\rho} W_\rho}_{\text{isotypics}}$$

The Fourier transform on a group is

$$\widehat{f}(\rho) = \sum_{x \in G} f(x)\, \rho(x) \qquad\qquad \rho \in \mathcal{R}$$

$$\widehat{f}(\rho) = \sum_{x \in G} f(x)\, \rho(x) \qquad f(x) = \frac{1}{|G|} \sum_{\rho \in \mathcal{R}} d_\rho \operatorname{tr}\left[\widehat{f}(\rho)\, \rho(x^{-1})\right]$$

1. Linearity: $\widehat{f+g} = \widehat{f} + \widehat{g}$
2. Unitarity: $\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle$
3. Left-translation: $\widehat{f^{z}(\rho)} = \rho(z)\, \widehat{f}(\rho)$
4. Convolution: $\widehat{f * g}(\rho) = \widehat{f}(\rho)\, \widehat{g}(\rho)$
5. The individual components correspond to different levels of smoothness.

What about the isotypics?

Recall the group algebra $\mathbb{C}G$:

$$f \cdot g = f * g \qquad\qquad f * g(x) = \sum_{y \in G} f(xy^{-1})\, g(y)$$

The isotypics are just the irreducible sub-algebras.

$$\mathbb{C}G = \bigoplus_{\rho \in \mathcal{R}} \mathrm{GL}(\mathbb{C}^{d_\rho})$$

isotypics

$$\mathbb{C}G = V_1 \oplus V_2 \oplus \ldots \oplus V_k$$

$$V_i = W_1 \oplus W_2 \oplus \ldots \oplus W_{d_{\rho_i}}$$

irreducible G-modules

Part 2

The symmetric group

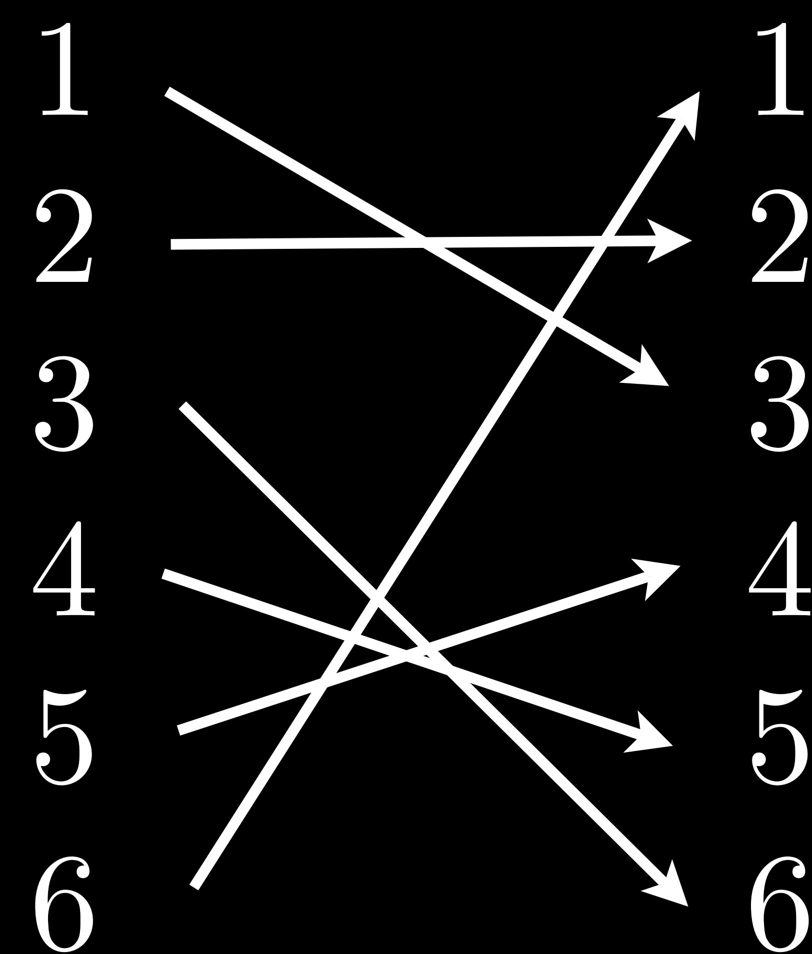
$\mathbb{S}_n$ is the group of bijections

$$\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$$

under composition of maps.

Clearly, $|\mathbb{S}_n| = n!$.

$$\sigma(1) = 3$$
$$\sigma(2) = 2$$
$$\sigma(3) = 6$$
$$\sigma(4) = 5$$
$$\sigma(5) = 4$$
$$\sigma(6) = 1$$



$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 6 & 5 & 4 & 1 \end{pmatrix}$$

Cycle notation

Example:

$$\sigma = (163)(45)(2)$$

Cycle type:

$$\sigma = (3, 2, 1)$$

Generators

Transpositions (i, j) generate the whole group.

In fact, adjacent transpositions are sufficient, since (assuming $i < j$)

$$(i, j) = (i, i+1) \ldots (j-2, -1)(j-1, j) \ldots (i\ +1, i+2)(i, i+1)$$

Subgroups

Cayley's theorem:

Any finite group G is a subgroup of $\mathbb{S}_{|\,G\,|}$.

Subgroups

$\mathbb{S}_k < \mathbb{S}_n$ permutes $\{1, 2, \ldots, k\}$

$\mathbb{S}_\lambda = \mathbb{S}_{\lambda_1} \times \mathbb{S}_{\lambda_2} \times \ldots \times \mathbb{S}_{\lambda_k} < \mathbb{S}_n$ permutes
$\{1, 2, \ldots, \lambda_1\}, \{\lambda_1 + 1, \ldots, \lambda_1 + \lambda_2\}, \ldots, \{n - \lambda_k, \ldots, n\}$

Normal subgroups

$$A_n = \{\, \sigma \in \mathbb{S}_n \mid \operatorname{sgn}(\sigma) = 1 \,\}$$

For $n \geq 5$ this is the only one!