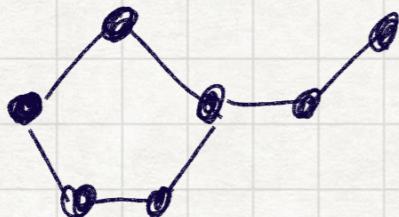


Representation of atomic structure , positions R
molecules



Solids



Tasks:

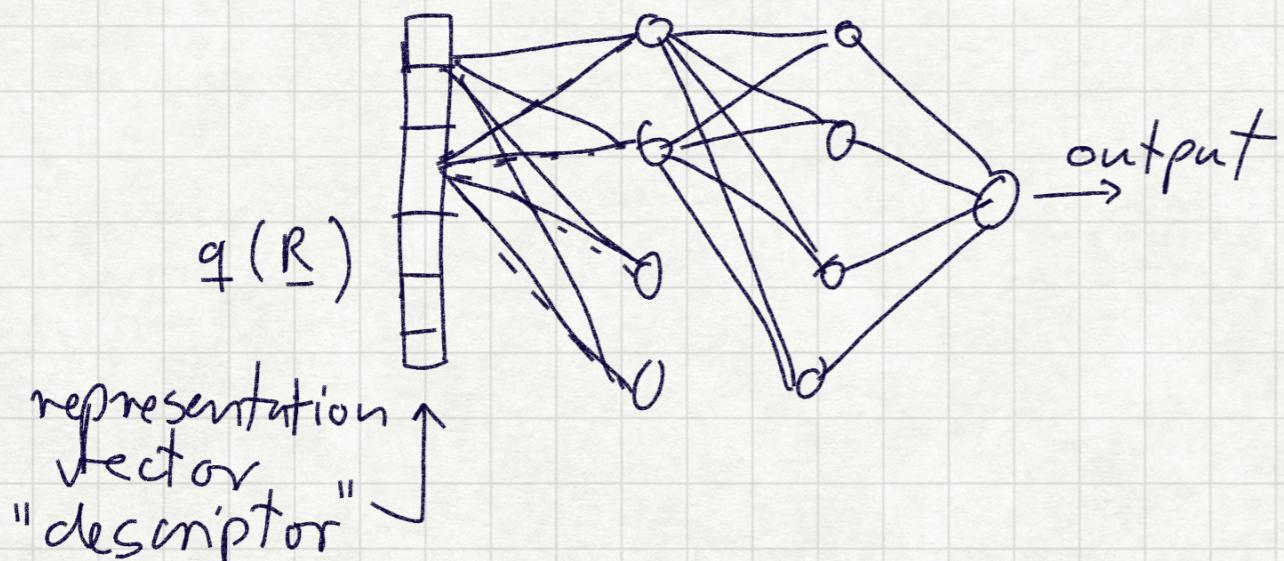
- Regression (force fields, other properties)
- Classification (discrete regression)
- Generative models : "chemical compound space"
"materials genome"

Symmetries:

translation, rotation, permutation

These representations go into ML models

e.g. neural network:



Kernel $K(q, q')$

$$f(\underline{R}) = F[q(\underline{R})] = \sum_i \alpha_i K(q, q_i)$$

Example:

Behler - Parrinello "Symmetry functions"

atom i $G_\alpha^1 = \sum_{j \neq i} e^{-\gamma_\alpha (R_{ij} - R_{\alpha}^s)} f_c(R_{ij})$ smooth cutoff function

$$G_\beta^2 = \sum_{j, k \neq i} (1 + \lambda_\beta \cos \theta_{ijk})^{\zeta_\beta} e^{-\gamma_\beta (R_{ij}^2 + R_{jk}^2 + R_{ik}^2)} f_c(R_{ij}) f_c(R_{jk}) f_c(R_{ik})$$

Invariants

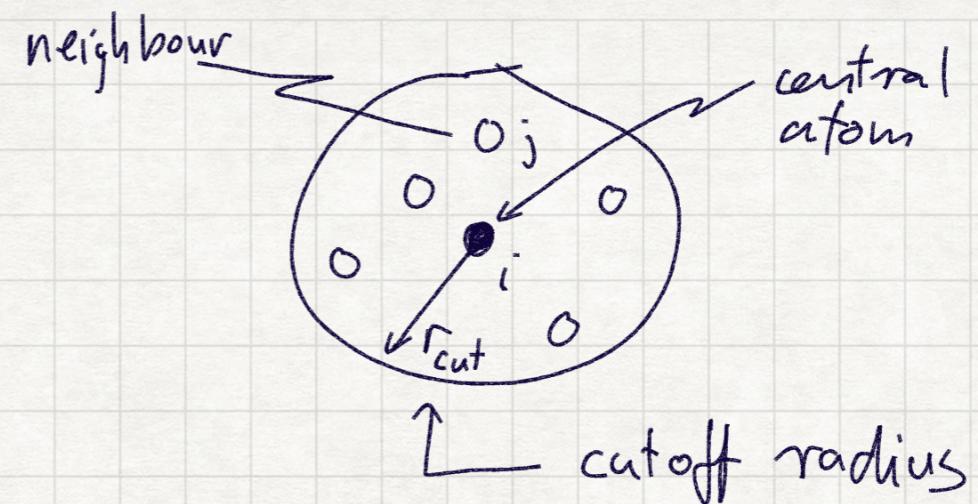
How to set parameters? Completeness?

Atomic neighbour density

$$g(\underline{r}) = \sum_j g(\underline{r} - \underline{R}_{ij}) f_c(R_{ij})$$

↑
convolution kernel : $\begin{cases} \delta(r) \\ e^{-|\underline{r}|^2/2\sigma^2} \end{cases}$

g is permutationally invariant



Rotational invariants

$$g(\underline{r}) = \sum_{nlm} c_{nlm} g_n(r) Y_{lm}(\hat{r})$$

Then $P_{nn'} = \sum_m c_{nlm}^* c_{nl'm}$ is rotationally invariant

Under mild assumptions P is complete.

Smooth overlap of
Atomic positions (SOAP)

$$\underline{f} \cdot \underline{f}' = \int d\hat{\underline{R}}^3 \left| \int d^3\underline{r} g(\underline{r}) g'(\hat{\underline{R}}\underline{r}) \right|^2$$

We can think of $P_{nm}^{\alpha\beta}$ as the projection of $\rho(\underline{r})$ onto basis functions:

$$b_\alpha(\underline{r}) = g_n(r) Y_{lm}(\hat{\underline{r}}) \quad \alpha: [n|lm]$$

But BP symmetry functions are also projections onto a (different) basis

$$b'_\alpha(\underline{r}) = e^{\gamma_\alpha(|\underline{r}| - R_\alpha^s)} f_c(|\underline{r}|)$$

[density kernel:]

$$G = \delta(\underline{r})$$

MBTR, von Lilienfeld's representations are also density projections, so are Mallat's scattering transforms

Which is the best basis to project onto?

Other cool properties of Yem basis

① Natural extension to basis functions to expand tensorial properties (Michele Ceriotti)

② Further invariants:

$$\sum_{m_1 m_2 m_3} \binom{\ell_1 \ell_2 \ell_3}{m_1 m_2 m_3} c_{n_1 \ell_1 m_1} c_{n_2 \ell_2 m_2} c_{n_3 \ell_3 m_3}$$

$$\sum_{m_1 m_2 m_3 m_4} \left[\begin{matrix} \ell_1 \ell_2 \ell_3 \ell_4 \\ m_1 m_2 m_3 m_4 \end{matrix} \right] c_{n_1 \ell_1 m_1} c_{n_2 \ell_2 m_2} c_{n_3 \ell_3 m_3} c_{n_4 \ell_4 m_4}$$

:

(Ralf Drautz)

③ Variant to describe entire structure, rather than neighbourhood

$$\bar{c}_{nlm} = \sum_i c_{nlm}^{(i)} \quad \leftarrow \text{average coefficients over atoms}$$

$$\bar{P}_{nn'l} = \sum_m \bar{c}_{nlm}^+ \bar{c}_{n'l'm}$$

(no cutoff!)