Oganov fingerprint feature

Malthe Kjær Bisbo

February 14, 2018

In the radial Oganov fingerprint feature the atomic radial distance below a cut-off radius R_c is separated into L bins by a set of sampling distances $r_1, r_2, ..., r_L$, where $r_l = (l-1)\Delta \quad , \quad l = 1, 2, ..., L$.

The feature is constructed by computing the gaussian smeared radial atom distribution for each atom and summing the contribution of each atom. Separate distribution vectors are computed for each atomic type combination (Z_A, Z_B). Thus for a structure containing Titanium and oxygen atoms, feature is created by concatenating the distribution vectors for O-O, O-Ti and Ti-Ti pairs. This is done in an ordered manner.

The contribution of atom i to the bin $[r_l, r_l + \Delta]$ of the (Z_i, Z) component of the feature is computed according to

$$g_{i,Z}(r_l) = \sum_{\substack{i \neq j \\ Z_j = Z}} \int_{r_l}^{r_l + \Delta} \frac{\delta(r - R_{ij})}{4\pi R_{ij}^2 \Delta \left(\frac{\gamma(Z_i, Z_j)}{V} \right)} dr \tag{1}$$

The full contribution of atom i to the (Z_i, Z) feature component is then.

$$\mathbf{g}_{i,Z} = \{g_{i,Z}(r_1), g_{i,Z}(r_2), ..., g_{i,Z}(r_L)\} \tag{2}$$

The total feature is then created by concatenation and summation over all atoms according to eq. 7 and 8 in Thomas' article.

Coming back to 1 δ in a Gaussian function with an appropriately choosen width. Considdering the denominator, $\gamma(Z_i, Z_j)$ is the number of $Z_i - Z_j$ pairs int he computational volume (unit cell) and V is the volume of the computational volume (volume of unit call). This means that $\left(\frac{\gamma(Z_i, Z_j)}{V} \right)$ is the average density of $Z_i - Z_j$ pairs in the computational volume.

Moving on, $4\pi R_{ij}^2 \Delta$ is the volume of the spherical shell at radius R_{ij} with thickness Δ. Thus the whole denominator represents the number of type pairs that is is expected in the feature bin corresponding to the spherical shell volume centered at R_{ij} if we assume that the atom distribution is uniform, which it certainly is not for any interesting structure.

Final remark:

The final feature of a structure, after summation but before adding $\mathbf{F}_0$ should thus be larger than one at atomic distances that are more common in the structure that they would be for a uniform distribution of atoms, and below one if they are less common. $\mathbf{F}_0 = -1$ thus seems like a reasonable choice.