

$$c) \phi(\vec{x}) = \begin{bmatrix} 1 \\ \vdots \\ \sqrt{3} x_i^2 \sqrt{\beta} \\ \vdots \\ \sqrt{6} x_i x_j \beta \\ \vdots \\ x_i^3 \sqrt{\beta^3} \\ \vdots \\ \sqrt{3} x_i^2 x_j \sqrt{\beta^3} \\ \vdots \\ \sqrt{6} x_i x_j x_k \sqrt{\beta^3} \\ \vdots \end{bmatrix}$$

The role of the parameter  $\beta$  is to scale each element of the feature vector  $\phi(\vec{x})$ . For smaller combinations, ~~the~~  $\beta$  is raised to a smaller power, and vice versa for larger combinations. ~~[1] [2] [3]~~

If  $0 < \beta < 1$ , it can be used to punish larger combinations in the feature vector.