- For $C > 2$, using maximal value method
- Augmented space
- $g_k\left(\underline{x}\right) = \underline{w}_k^{\ T}\ \underline{x}, \quad k = 1,2,\cdots,C$

Decision rule:

$$g_k\left(\underline{x}\right) > g_j\left(\underline{x}\right)\ \ \forall j \neq k\ \ \Rightarrow\ \ \underline{x} \in S_k$$

Algorithm:

1. Shuffle the order of training data points
2. For each data point $\underline{x}^{(k)}$ [Given $\underline{x}^{(k)} \in S_k$]:

   If $g_k\left(\underline{x}^{(k)}\right) > g_j\left(\underline{x}^{(k)}\right)\ \ \forall j \neq k$

   then: $\underline{w}^{(m)}\left(i+1\right) = \underline{w}^{(m)}\left(i\right)\ \ \forall m$

   else:

   $$\underline{w}^{(k)}\left(i+1\right) = \underline{w}^{(k)}\left(i\right) + \eta\left(i\right)\underline{x}^{(k)}$$

   Let $l = \arg\max_{j \neq k}\left\{g_j\left(\underline{x}^{(k)}\right)\right\}$

   (If more than one possible value of l, pick any one)

   $$\underline{w}^{(l)}\left(i+1\right) = \underline{w}^{(l)}\left(i\right) - \eta\left(i\right)\underline{x}^{(k)}$$

   $$\underline{w}^{(m)}\left(i+1\right) = \underline{w}^{(m)}\left(i\right)\ \ \forall m \neq l,k$$

   until: all training data points are correctly classified.

Convergence: Convergence is proven in DHS 5.12.2 for $\eta\left(i\right) = \text{constant} > 0$, for linearly separable data.