$$egin{aligned} oldsymbol{W}_f^T oldsymbol{W}_{t+1} &= oldsymbol{W}_f^T oldsymbol{W}_t + y_n(t) oldsymbol{X}_n(t) \ &\geqslant oldsymbol{W}_f^T oldsymbol{W}_t + \min_n y_n(t) oldsymbol{W}_f^T oldsymbol{X}_n(t) \end{aligned}$$

且有 $W_0 = 0$,故有 $\boldsymbol{W}_f^T \boldsymbol{W}_T \geqslant T \cdot \min_n y_n \boldsymbol{W}_f^T \boldsymbol{X}_n$;

又由于

$$\|\boldsymbol{W}_{t+1}\|^{2} = \|\boldsymbol{W}_{t} + y_{n}(t)\boldsymbol{X}_{n}(t)\|^{2}$$

$$= \|\boldsymbol{W}_{t}\|^{2} + 2y_{n}(t)\boldsymbol{W}_{t}^{T}\boldsymbol{X}_{n}(t) + \|y_{n}(t)\boldsymbol{X}_{n}(t)\|^{2}$$

$$\leq \|\boldsymbol{W}_{t}\|^{2} + 0 + \|y_{n}(t)\boldsymbol{X}_{n}(t)\|^{2}$$

$$\leq \|\boldsymbol{W}_{t}\|^{2} + \max \|\boldsymbol{X}_{n}(t)\|^{2}$$

故有 $\|\boldsymbol{W}_T\| \leqslant \sqrt{T \cdot \max_n \|\boldsymbol{X}_n\|^2}$;

综上所述,有

$$egin{aligned} egin{aligned} oldsymbol{W}_f^T oldsymbol{W}_T & oldsymbol{X}_n \ \|oldsymbol{W}_f^T oldsymbol{W}_T & oldsymbol{W}_T oldsymbol{X}_n \ \|oldsymbol{W}_f \| \cdot \sqrt{T \cdot \max_n \|oldsymbol{X}_n\|^2} \ &= \sqrt{T} \cdot constant \end{aligned}$$

4,针对线性可分训练样本集,PLA 算法中,假设对分错样本进行了 T 次纠正后得到的分类面不再出现错分状况,定义: $R^2 = \max ||x_n||^2$,

$$\rho = \min_{n} y_n \frac{\mathbf{w}_f^T}{\|\mathbf{W}_f\|} \mathbf{x}_n, \quad$$
试证明: $\mathbf{T} \leq \frac{\mathbf{R}^2}{\rho^2}$

证明:

$$egin{aligned} & rac{oldsymbol{W}_f^T oldsymbol{W}_T}{\|oldsymbol{W}_f^T \| oldsymbol{W}_T \| oldsymbol{W}_T \| oldsymbol{W}_T \| oldsymbol{W}_T \| oldsymbol{W}_T \| oldsymbol{W}_T \| oldsymbol{V}_T & \min_n y_n oldsymbol{W}_f^T oldsymbol{X}_n \ & \| oldsymbol{W}_f \| oldsymbol{W}_T \| oldsymbol{V}_T & \max_n \| oldsymbol{X}_n \|^2 \ & = \sqrt{T} \cdot rac{
ho}{R} \end{aligned}$$