公式(7)和(8)的推导，推导过程中请注意 a_k 与 b_k 在窗口 w_k 中是恒定的：

$$
\begin{aligned}
0=\frac{\partial E}{\partial b_k} &= \sum_{i\in w_k} 2(a_k I_i + b_k - p_i) \\
b_k &= \sum_{i\in w_k} (p_i - a_k I_i) \\
&= \sum_{i\in w_k} p_i - a_k \sum_{i\in w_k} I_i \\
&= \bar{p}_k - a_k \mu_k
\end{aligned}
\tag{1-1}
$$

$$
\begin{aligned}
0=\frac{\partial E}{\partial a_k} &= \sum_{i\in w_k} (2(a_k I_i + b_k - p_i) I_i + 2\varepsilon a_k) \\
&= \sum_{i\in w_k} ((a_k I_i + b_k - p_i) I_i + \varepsilon a_k) \\
&= \sum_{i\in w_k} ((a_k I_i + \bar{p}_k - a_k\mu_k - p_i) I_i + \varepsilon a_k) \\
&= \sum_{i\in w_k} ((a_k I_i^2 + \varepsilon a_k - a_k \mu_k I_i) + (\bar{p}_k - p_i) I_i) \\
a_k &= \frac{\sum_{i\in w_k} (p_i - \bar{p}_k) I_i}{\sum_{i\in w_k} (I_i^2 + \varepsilon - \mu_k I_i)} \\
&= \frac{\sum_{i\in w_k} p_i I_i - \bar{p}_k \sum_{i\in w_k} I_i}{\sum_{i\in w_k} (I_i^2 + \varepsilon - \mu_k I_i)} \\
&= \frac{\frac{1}{|w|}\sum_{i\in w_k} p_i I_i - \frac{1}{|w|}\bar{p}_k \sum_{i\in w_k} I_i}{\frac{1}{|w|}\sum_{i\in w_k} (I_i^2 + \varepsilon - \mu_k I_i)} \\
&= \frac{\frac{1}{|w|}\sum_{i\in w_k} p_i I_i - \bar{p}_k \mu_k}{\frac{1}{|w|}\sum_{i\in w_k} (I_i^2 - \mu_k I_i) + \varepsilon} \\
&= \frac{\frac{1}{|w|}\sum_{i\in w_k} p_i I_i - \bar{p}_k \mu_k}{\frac{1}{|w|}\sum_{i\in w_k} I_i^2 - \frac{1}{|w|}\mu_k \sum_{i\in w_k} I_i + \varepsilon} \\
&= \frac{\frac{1}{|w|}\sum_{i\in w_k} p_i I_i - \bar{p}_k \mu_k}{\frac{1}{|w|}\sum_{i\in w_k} I_i^2 - \mu_k^2 + \varepsilon} \\
&= \frac{\frac{1}{|w|}\sum_{i\in w_k} p_i I_i - \bar{p}_k \mu_k}{\delta_k^2 + \varepsilon}
\end{aligned}
\tag{1-2}
$$

式(1-2)的计算借助了概率公式：

$$E[(X-E(X))^2] = E(X^2) - E^2(X)$$

公式(19)和(20)的推导，推导过程中请注意 $\boldsymbol{a}_k$ 与 b_k 在窗口 w_k 中是恒定的：

$$
\begin{aligned}
0 = \frac{\partial E}{\partial b_k} &= \sum\nolimits_{i \in w_k} 2(\mathbf{a}_k^T \mathbf{I}_i + b_k - p_i) \\
b_k &= \sum\nolimits_{i \in w_k} (p_i - \mathbf{a}_k^T \mathbf{I}_i) \\
&= \sum\nolimits_{i \in w_k} p_i - \mathbf{a}_k^T \sum\nolimits_{i \in w_k} \mathbf{I}_i \\
&= \bar{p}_k - \mathbf{a}_k^T \boldsymbol{\mu}_k
\end{aligned} \tag{1-3}
$$

$$
\begin{aligned}
0 = \frac{\partial E}{\partial \mathbf{a}_k} &= \sum\nolimits_{i \in w_k} (2(\mathbf{a}_k^T \mathbf{I}_i + b_k - p_i)\mathbf{I}_i + 2\varepsilon \mathbf{a}_k) \\
&= \sum\nolimits_{i \in w_k} ((\mathbf{a}_k^T \mathbf{I}_i + b_k - p_i)\mathbf{I}_i + \varepsilon \mathbf{a}_k) \\
&= \sum\nolimits_{i \in w_k} ((\mathbf{a}_k^T \mathbf{I}_i + \bar{p}_k - \mathbf{a}_k^T \boldsymbol{\mu}_k - p_i)\mathbf{I}_i + \varepsilon \mathbf{a}_k) \\
&= \sum\nolimits_{i \in w_k} ((\mathbf{a}_k^T \mathbf{I}_i \mathbf{I}_i + \varepsilon \mathbf{a}_k - \mathbf{a}_k^T \boldsymbol{\mu}_k \mathbf{I}_i) + (\bar{p}_k - p_i)\mathbf{I}_i) \\
&= \sum\nolimits_{i \in w_k} ((\mathbf{I}_i^T \mathbf{a}_k \mathbf{I}_i + \varepsilon \mathbf{U} \mathbf{a}_k - \boldsymbol{\mu}_k^T \mathbf{a}_k \mathbf{I}_i) + (\bar{p}_k - p_i)\mathbf{I}_i) \\
&= \sum\nolimits_{i \in w_k} ((\mathbf{I}_i^T \mathbf{I}_i \mathbf{a}_k + \varepsilon \mathbf{U} \mathbf{a}_k - \boldsymbol{\mu}_k^T \mathbf{I}_i \mathbf{a}_k) + (\bar{p}_k - p_i)\mathbf{I}_i) \\
\mathbf{a}_k &= \frac{\sum\nolimits_{i \in w_k} (p_i - \bar{p}_k)\mathbf{I}_i}{\sum\nolimits_{i \in w_k} (\mathbf{I}_i^T \mathbf{I}_i + \varepsilon \mathbf{U} - \boldsymbol{\mu}_k^T \mathbf{I}_i)} \\
&= \frac{\sum\nolimits_{i \in w_k} p_i \mathbf{I}_i - \bar{p}_k \sum\nolimits_{i \in w_k} \mathbf{I}_i}{\sum\nolimits_{i \in w_k} (\mathbf{I}_i^T \mathbf{I}_i + \varepsilon \mathbf{U} - \boldsymbol{\mu}_k^T \mathbf{I}_i)} \\
&= \frac{\frac{1}{|w|} \sum\nolimits_{i \in w_k} p_i \mathbf{I}_i - \frac{1}{|w|} \bar{p}_k \sum\nolimits_{i \in w_k} \mathbf{I}_i}{\frac{1}{|w|} \sum\nolimits_{i \in w_k} (\mathbf{I}_i^T \mathbf{I}_i + \varepsilon \mathbf{U} - \boldsymbol{\mu}_k^T \mathbf{I}_i)} \\
&= \frac{\frac{1}{|w|} \sum\nolimits_{i \in w_k} p_i \mathbf{I}_i - \bar{p}_k \boldsymbol{\mu}_k}{\frac{1}{|w|} \sum\nolimits_{i \in w_k} (\mathbf{I}_i^T \mathbf{I}_i - \boldsymbol{\mu}_k^T \mathbf{I}_i) + \varepsilon \mathbf{U}} \\
&= \frac{\frac{1}{|w|} \sum\nolimits_{i \in w_k} p_i \mathbf{I}_i - \bar{p}_k \boldsymbol{\mu}_k}{\frac{1}{|w|} \sum\nolimits_{i \in w_k} \mathbf{I}_i^T \mathbf{I}_i - \frac{1}{|w|} \boldsymbol{\mu}_k^T \sum\nolimits_{i \in w_k} \mathbf{I}_i + \varepsilon \mathbf{U}} \\
&= \frac{\frac{1}{|w|} \sum\nolimits_{i \in w_k} p_i \mathbf{I}_i - \bar{p}_k \boldsymbol{\mu}_k}{\frac{1}{|w|} \sum\nolimits_{i \in w_k} \mathbf{I}_i^T \mathbf{I}_i - \boldsymbol{\mu}_k^T \boldsymbol{\mu}_k + \varepsilon \mathbf{U}} \\
&= \frac{\frac{1}{|w|} \sum\nolimits_{i \in w_k} p_i \mathbf{I}_i - \bar{p}_k \boldsymbol{\mu}_k}{\boldsymbol{\Sigma}_k + \varepsilon \mathbf{U}}
\end{aligned} \tag{1-4}
$$

式(1-4)的计算借助了概率公式：

$$E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y)$$

备注：

粗体表示向量和矩阵。