e. 对于多分类.

$S_t = S_b + S_w = \sum_{i=1}^{m} (x_i - \mu)(x_i - \mu)^T$

其中 μ 是所有样例均值向量，N个类

$S_w = \sum_{i=1}^{N} S_{w_i}$ ， $S_{w_i} = \sum_{x \in X_i} (x - \mu_i)(x - \mu_i)^T$

$S_b = S_t - S_w = \sum_{i=1}^{N} m_i (\mu_i - \mu)(\mu_i - \mu)^T$.

优化目标 $\max_W \frac{tr(W^T S_b W)}{tr(W^T S_w W)}$ $W \in R^{d \times (N-1)}$

通过 $S_b W = \lambda S_w W \rightarrow$ d' 个最大非零广义特征值. $d' \leq N-1$

从而 W 为投影矩阵，多分类 LDA 投影到 d' 维，$< d$.

降维.