we'll take derivative from $\underline{a}$ $\theta_{C,k}$ which pertains to $\underline{a}$ C
this will make the derivative easier to $\underline{\text{see}}$.

$$L(P(X,y)) = \sum_{i=1}^{m} \{ \log(P(y^i=c)) + \sum_{k=1}^{n} \log \theta_{c,k}^{x^i_k} \} + \lambda(\sum_{k=1}^{n} \theta_{c,k} - 1)$$

$$\frac{\partial (P(X,y))}{\partial \theta_{c,k}} = \sum_{i=1}^{m} \{ 0 + \frac{1}{\theta_{c,k}} 1^{x^i_{ck}} \} + \lambda = 0$$

Because we are differentiating for a particular $\theta_{c,k}$ hence only one term from $\lambda \sum_{k=1}^{n} \theta_{c,k}$ can be differenciated to 1. the rest are constants.

$1^{x^i_{ck}}$: indicates the presence of word k in group c in sentence i
so if the word k exist in sentence i group c the value is $\underline{1}$ otherwise it is $\underline{0}$.

$$\frac{\partial (P(X,y))}{\partial \theta_{c,k}} = \frac{1}{\theta_{c,k}} \sum_{i=1}^{m} 1^{x^i_{ck}} + \lambda = 0 \Rightarrow \theta_{c,k} = \frac{-\sum_{i=1}^{m} 1^{x^i_{ck}}}{\lambda} \quad (*)$$

$$\text{since } \sum_{k=1}^{n} \theta_{c,k} = 1 \Rightarrow \sum_{k=1}^{n} \theta_{c,k} = \frac{-\sum_{k=1}^{n}\sum_{i=1}^{M} 1^{x^i_{ck}}}{\lambda} = 1$$

$$\Rightarrow \lambda = -\sum_{k=1}^{n}\sum_{i=1}^{m} 1^{x^i_{ck}}$$

$\Rightarrow$ replace λ in (*)