(Multivariate) Sum of Squared Differences from the mean

MLE for the Normal Gaussian

$$\sum_{i=0}^{N-1} (x^{[i]}-\mu)^2 \longrightarrow \sum_{i=0}^{N-1} (x^{[i]}-\bar{x})^2 + N(\bar{x}-\mu)^2$$

Multivariate: $\sum_{i=0}^{N-1} \left((\underline{x}^{[i]}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{x}^{[i]}-\underline{\mu}) \right)$

Univariate

$$\sum_{i=0}^{N-1} (x^{[i]}-\mu)^2 = \sum_{i=0}^{N-1} (x^{[i]}-\bar{x})^2 + N\cdot(\bar{x}-\mu)^2$$

$$\bar{x} = \frac{1}{N}\sum_{i=0}^{N-1} x^{[i]} = \mu_{MLE}$$

$$\Leftrightarrow \sum_{i=0}^{N-1} x^{[i]} = N\cdot\bar{x}$$

$$\sum_{i=0}^{N-1} (x^{[i]}-\mu)^2 = \sum_{i=0}^{N-1} \left(x^{[i]^2} - 2x^{[i]}\mu + \mu^2 \right)$$

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2}\right) - 2\mu \underbrace{\sum_{i=0}^{N-1} x^{[i]}}_{N\bar{x}} + N\mu^2$$

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2}\right) + N\left(\mu^2 - 2\mu\bar{x}\right) \quad + \bar{x}^2 - \bar{x}^2$$

Complete the square

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2}\right) + N\left(\underbrace{\mu^2 - 2\mu\bar{x} + \bar{x}^2}_{(\mu-\bar{x})^2} - \bar{x}^2\right)$$

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2}\right) + N(\mu-\bar{x})^2 - N\bar{x}^2$$

$$+N\bar{x}^2 - N\bar{x}^2$$

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2}\right) - 2\underbrace{N\bar{x}}_{\sum_{i=0}^{N-1} x^{[i]}}\cdot\bar{x} + \underbrace{N\bar{x}^2}_{\sum_{i=0}^{N-1} \bar{x}^2} + N(\mu-\bar{x})^2$$

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2}\right) - 2\bar{x}\sum_{i=0}^{N-1} x^{[i]} + \sum_{i=0}^{N-1} \bar{x}^2 + N(\mu-\bar{x})^2$$

$$= \sum_{i=0}^{N-1} \left(x^{[i]^2} - 2\bar{x}x^{[i]} + \bar{x}^2 \right) + N(\mu-\bar{x})^2$$

$$= \sum_{i=0}^{N-1} (x^{[i]}-\bar{x})^2 + N(\mu-\bar{x})^2 \quad \square$$

why? → μ is no longer in $\sum_{i=0}^{N-1}$

Multivariate:

$$\sum_{i=0}^{N-1} \left((\underline{x}^{[i]}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{x}^{[i]}-\underline{\mu}) \right)$$

$\underline{\underline{\Sigma}}^{-1}$

$$= \sum_{i=0}^{N-1} \left((\underline{x}^{[i]}-\underline{\bar{x}})^T \underline{\underline{\Lambda}} (\underline{x}^{[i]}-\underline{\bar{x}}) \right) + N(\underline{\bar{x}}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{\bar{x}}-\underline{\mu})$$

$$\underline{\bar{x}} = \frac{1}{N}\sum_{i=0}^{N-1} \underline{x}^{[i]} = \underline{\mu}_{MLE}$$

$$\Leftrightarrow \sum_{i=0}^{N-1} \underline{x}^{[i]} = N\cdot\underline{\bar{x}}$$

$$\sum_{i=0}^{N-1} \left((\underline{x}^{[i]}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{x}^{[i]}-\underline{\mu}) \right)$$

$$= \sum_{i=0}^{N-1} \left(\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{x}^{[i]} - 2\underline{\mu}^T \underline{\underline{\Lambda}}\, \underline{x}^{[i]} + \underline{\mu}^T \underline{\underline{\Lambda}}\, \underline{\mu} \right)$$

$$= \sum_{i=0}^{N-1} \left(\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{x}^{[i]} \right) - 2\underline{\mu}^T \underline{\underline{\Lambda}} \underbrace{\left(\sum_{i=0}^{N-1} \underline{x}^{[i]} \right)}_{N\underline{\bar{x}}} + N\underline{\mu}^T \underline{\underline{\Lambda}}\, \underline{\mu}$$

$$= \sum_{i=0}^{N-1} \left(\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{x}^{[i]} \right) + N\left(\underline{\mu}^T \underline{\underline{\Lambda}}\, \underline{\mu} - 2\underline{\mu}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}} \qquad \right)$$

$$\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}} - \underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}}$$

$$= \sum_{i=0}^{N-1} \left(\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{x}^{[i]} \right) - N\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}} + N(\underline{\bar{x}}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{\bar{x}}-\underline{\mu})$$

$$+N\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}} - N\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}}$$

Sym

$$= \sum_{i=0}^{N-1} \left(\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{x}^{[i]} \right) - 2N\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}} + \underbrace{N\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}}}_{\sum_{i=0}^{N-1} (\underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}})} + N(\underline{\bar{x}}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{\bar{x}}-\underline{\mu})$$

$$\sum_{i=0}^{N-1} \underline{x}^{[i]}$$

$$= \sum_{i=0}^{N-1} \left(\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{x}^{[i]} - 2\underline{x}^{[i]^T} \underline{\underline{\Lambda}}\, \underline{\bar{x}} + \underline{\bar{x}}^T \underline{\underline{\Lambda}}\, \underline{\bar{x}} \right)$$

$$+ N(\underline{\bar{x}}-\underline{\mu}) \underline{\underline{\Lambda}} (\underline{\bar{x}}-\underline{\mu})$$

$$= \sum_{i=0}^{N-1} \left((\underline{x}^{[i]}-\underline{\bar{x}})^T \underline{\underline{\Lambda}} (\underline{x}^{[i]}-\underline{\bar{x}}) \right) + N(\underline{\bar{x}}-\underline{\mu})^T \underline{\underline{\Lambda}} (\underline{\bar{x}}-\underline{\mu})$$

$\square$