Ass5 Solutions

Chang Li

October 2, 2016

1 Solutions

1.1 Linear Programming

1.2 Regularized Maximum Likelihood Estimation

1.2.1 a

Because dataset D is drawn from a Gaussian distribution with mean μ and covariance Σ. Suppose $x \in R^n$, so the density is:

$$p(x) = (2\pi)^{-n/2} det(\Sigma)^{-1/2} exp(-(x-\mu)^T \Sigma^{-1}(x-\mu)/2)$$

The average log-likelihood function has the form:

$$\begin{aligned} l(D) &= E(\log p(x_1, x_2, \dots, x_m)) \\ &= E(-\log\,(2\pi) - \log\, det\Sigma - \sum_{i=1}^{M}(-(x_i-\mu)^T\Sigma^{-1}(x_i-\mu))) \end{aligned}$$

Since $(x-\mu)^T\Sigma^{-1}(x-\mu)$ is a number and trace(AB) = trace(BA) therefore,

$$(x-\mu)^T\Sigma^{-1}(x-\mu) = \mathrm{tr}((x-\mu)^T\Sigma^{-1}(x-\mu)) = \mathrm{tr}(\Sigma^{-1}(x-\mu)^T(x-\mu))$$

Plugging into expectation we get:

$$\begin{aligned} E[(x-\mu)^T\Sigma^{-1}(x-\mu)] &= E[\mathrm{tr}((x-\mu)^T(x-\mu)\Sigma^{-1}])] \\ &= \mathrm{tr}(E[(x-\mu)^T(x-\mu)\Sigma^{-1}]) \\ &= \mathrm{tr}(\hat{\Sigma}\Sigma^{-1}) \end{aligned}$$

Plugging into the third term of $l(D)$ we can reformulate it as:

$$\begin{aligned} l(D) &= -\log\, det\Sigma - tr(\hat{\Sigma}\Sigma^{-1}) + const \\ &= \log\, det\Sigma^{-1} - tr(\hat{\Sigma}\Sigma^{-1}) + const \end{aligned}$$

1.2.2 b

Because $K \succ 0$, we use the fact in section 3.1.5 [1] that the log of determinant is concave so its negation is convex and tr$(\hat{\Sigma}K)$ is also convex. And the absolute value function and summation operation are also convex operations. Therefore the regularized maximum likelihood optimization problem is a summation of convex functions and henceTotal Variation Denoising convex.

images/Q1_code.png

Figure 1: "fawe"

1.2.3 c

1.3 Total Variation Denoising

For this question we are looking for a convex optimization problem. You do not need to give a closed-form solution.

Bibliography

[1] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.