Homework 4 Haocong Wang mw814

1. $f(x_1, x_2) = e^{x_1+3x_2-0.1} + e^{x_1-3x_2-0.1} + e^{-x_1-0.1}$

$$\Rightarrow \begin{cases} \nabla_{x_1} f = e^{x_1+3x_2-0.1} + e^{x_1-3x_2-0.1} - e^{-x_1-0.1} \\ \nabla_{x_2} f = 3e^{x_1+3x_2-0.1} - 3e^{x_1-3x_2-0.1} \end{cases}$$

Use backtracking line search and set $\alpha = 0.1$, $\beta = 0.6$.

With t starting from in each step, we repeat the process

If $f(x+td) < f(x) + \alpha t \langle d$, converged ; else, $t = \beta t$

$x_1 = x_1 - t\nabla_{x_1} f(x)$; $x_2 = x_2 - t\nabla_{x_2} f(x)$.

2. With Newton method, $\nabla f = \alpha' \cdot \frac{1}{1-\alpha x} - \frac{1}{1+x} + \frac{1}{1-x}$. $H = \nabla^2 f(x)$

Use the same method in Q1 to do the gradient descent.

3. $\min_{w,b} \|w\|_2^2$ s.t. $y_i(b - \langle x_i, w\rangle) + 1 \le 0 \Rightarrow \min_{w,b} \ell(y_i(b - \langle x_i, w\rangle) + 1) + \|w\|_2^2$

$$\ell(u) = (u)_+ = \begin{cases} 0, & u \le 0 \\ u, & u > 0 \end{cases}$$

For SVM, we have

$$y_i(b - \langle x_i, w\rangle) + 1 = [-y_i x_i, y_i] \begin{bmatrix} w \\ b \end{bmatrix} + 1$$

$$\Rightarrow A_i = \begin{bmatrix} -y_1 x_1 & y_1 \\ \vdots & \\ -y_{ni} x_{ni} & y_{ni} \end{bmatrix}, \quad V = \begin{bmatrix} w \\ b \end{bmatrix} \in \mathbb{R}^{n+1}. \quad V(u) = \sum_{i=1}^{n} |V[n]|^2$$

$$\Rightarrow \min_{V} \sum_{i=1}^{n} l_i(A_i V + 1) + r(v) \Rightarrow \min_{V_1 \cdots V_n} \sum_{i=1}^{n} l_i(A_i V_i + 1) + r(z), \quad \text{s.t. } V_i - z = 0$$

The argumented Lagrangian : $L_\rho(V_1, \cdots V_n, z, \mu_1, \cdots, \mu_n)$

$$= \sum_{i=1}^{n} l_i(A_i x + 1) + \frac{\rho}{2} \sum_{i=1}^{n} \|V_i - z + \mu_i\|_2^2 + r(z)$$

So ADMM iteration : $V_i^{(k+1)} = \arg\min_{V_i} \left(l_i(A_i V_i + 1) + \frac{\rho}{2} \|V_i - z^{(k)} + \mu_i^{(k)}\|_2^2 \right)$

$$z^{(k+1)} = \arg\min_{z} \left(r(z) + \frac{\rho}{2} \|z - x_i^{(k+1)} - \mu_i^{(k)}\|_2^2 \right)$$

$$\mu_i^{(k+1)} = \mu_i^{(k)} + x_i^{(k+1)} - z^{(k+1)} \qquad i = 1, \cdots, n$$

$$\Rightarrow \begin{cases} V_i^{(k+1)} = \arg\min \left(L'(A_i V_i + 1) + \frac{\rho}{2} \|V_i - z^{(k)} + \mu_i^{(k)}\|_2^2 \right) \\ z_{1:N}^{(k+1)} = \frac{\rho}{1+N\rho} \left(\overline{V}_{1:N}^{(k+1)} + \overline{\mu}_{1:N}^{(k)} \right), \quad z_{N+1}^{(k+1)} = \overline{V}_{N+1}^{(k+1)} + \overline{\mu}_{N+1}^{(k)} \\ \mu_i^{(k+1)} = \mu_i^{(k)} + V_i^{(k+1)} - z^{(k+1)}, \quad i = 1, \cdots, N \end{cases}$$