Assignment 2

1. a. $\mathbb{1}(z)_\chi = \max_{z \in \mathbb{R}^n} \{ z^T x - \mathbb{1}(x)_\chi \}$

$= \max_{z \in \chi} \{ z^T x \} = S(x)$

~~$S^*(x) = \mathbb{1}(x)_\chi$ b/c The conjugate of a conjugate is itself for closed, convex, proper functions~~

$S^*(x) = \mathbb{1}(x)_\chi$ b/c The conjugate of a conjugate is itself for closed, convex, proper functions

b. ~~$\max_{z \in \mathbb{R}^n} \{ z^T x - \|x\| \}$~~

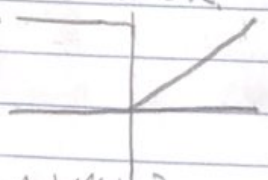
$f(x) = \|x\| = \max_{\|z\|_2 \le 1} \{ z^T x \} = \max_{z \in \chi} \{ z^T x \} = S(x)$

So $f^*(x) = \mathbb{1}(x)$ from part A.

c. $\partial f(x) = \operatorname{argmin}_{z \in \mathbb{R}^n} f^*(x) = \operatorname{argmin}_{x \in \mathbb{R}^n} \mathbb{1}(x)_\chi = \chi$

0 when $x \in \chi$

2 a. $\psi(x) = \begin{cases} x & x \ge 0 \\ \infty & x < 0 \end{cases}$



a. $S_\lambda(y) = \operatorname{prox}_{\lambda\psi}(y) = \operatorname{argmin}_x \left\{ \frac{1}{2}(x-y)^2 + \lambda \psi(x) \right\}$

$= \begin{cases} \frac{1}{2}(x-y)^2 + \lambda x & x \ge 0 \\ \infty & x < 0 \end{cases}$

Note no min for ∞ So focus on $x \ge 0$.

Let $x \geq 0$:

$\operatorname{argmin} \left\{ \frac{1}{2}(x-y)^2 + \lambda x \right\} \Rightarrow x - y + \lambda = 0$

$x^* = y - \lambda$ for $y > \lambda$ to make x stay > 0.

Let $x = 0$. $\rightarrow$ no ∇ so need g^T for λx.

$\operatorname{argmin} \Rightarrow x - y + \lambda g^T = 0$

$x^* = y - \lambda g^T = 0 \quad y \leq \lambda$

$\phi(y) \geq \phi(x) + g^T(y - x)$

since $x = 0$.

$\frac{\phi(y)}{y} \geq g^T$

$\frac{y}{y} = 1 \geq g^T$

$-\infty < g^T < 1$

$$\operatorname{prox}_{\lambda g}(y) = \begin{cases} y - \lambda & y > \lambda \\ 0 & y < \lambda \end{cases}$$

b. $\operatorname{prox}_{\lambda \|x\|}(y) = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2}\|x - y\|^2 + \lambda \|x\| \right\} = \begin{cases} \left(\|y\| - \lambda\right)^+ \frac{y}{\|y\|} & y \neq 0 \\ 0 & y = 0 \end{cases}$

$= \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 + \|x\|\|y\| + \lambda \|x\| \right\}$

$= ?$

3. $x^+ = \text{prox}_{\gamma g}(x^{k-1}) = \text{argmin}\left\{ \frac{1}{2}\|x - x^{k-1}\|^2 + \gamma g(x) \right\}$

a. $x^+ = \text{prox}_{\gamma g}(x) = x - \gamma g_\gamma(x)$ w/ $g_\gamma(x) = \frac{1}{\gamma}(x - \text{prox}_{\gamma g}(x))$.

$x \in \mathbb{R}^n$ $\gamma > 0$ $g \in \Gamma^0(\mathbb{R}^n)$

by Theorem 3 of Lec 15 since $g \in \Gamma^0(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$

$x = \text{prox}_g(z) \iff (z - x) \in \partial g(z) \iff (z-x)^T(y-x) \le f(y) - f(x)$.

where $z = x^+$

$x^+ - x \in \partial g(x^+) \longrightarrow -\gamma g_\gamma(x) \in \partial g(x^+)$

$\text{argmin}\left\{ \frac{1}{2}\|x^+ - x\|^2 + \gamma g(x) \right\} = x^+ - x + \gamma \partial g(x) = 0$

$x^+ - x = -\gamma g(x) \longrightarrow -\gamma g_\gamma(x) \in -\gamma \partial g(x)$

$g_\gamma(x) \in \partial g(x)$ ✓

b. $g(x^+) \le g_\gamma(x) + \nabla g_\gamma(x)^T(x^+ - x) + \frac{L}{2}\|x^+ - x\|_2^2$

$\le g_\gamma(x) + g_\gamma(x)^T(x^+ - x) + \frac{\gamma}{2}\|g_\gamma(x)\|_2^2$ (pars; $\gamma \le 1/L$)

$= g(x) + g_\gamma(x)^T(x^+ - z) - \frac{\gamma}{2}\|g_\gamma(x)\|_2^2$

$g(x^+) \le g_\gamma(x) + g_\gamma(x)^T(x^+ - x) - \frac{\gamma}{2}\|g_\gamma(x)\|_2^2$

Looking at the inequality we know $\|g_\gamma(x)\|_2^2 > 0$ so subtracting that will give a lower value, else $g^T(x)$ is a descent direction so $g(x^+) \le g(x)$. ✓

c. $g \in \Gamma^0(\mathbb{R}^n)$ $\quad g^* = g(x^*)$ $\quad \gamma_k = 1/\sqrt{k}$

$\frac{r}{2} \sum_1^k \|g(x^{k-1})\|_2^2 \le g(x^0) - g(x^1) + g(x^1) - g(x^2) + \dots = g(x^0) - g(x^*)$

where $g(x^*) \le g(x^k)$

$g(x^k) - g(x^*) \le \frac{1}{2r} \left[\|x^{k-1} - x^*\|_2^2 - \|x^k - x^*\|_2^2 \right]$

$g(x^k) - g(x^*) \le \frac{1}{2r} \left[\|x^0 - x^*\|_2^2 - \|x^k - x^*\|_2^2 \right]$

$g(x^k) - g(x^*) \le \frac{1}{2r} \|x^0 - x^*\|_2^2$ $\quad$ Let $r = 1/\sqrt{k}$

$g(x^k) - g(x^*) \le \frac{1}{2\sqrt{k}} \|x^0 - x^*\|_2^2$

d. if $\gamma_k = \gamma > 0$ then $g(x^k) - g(x^*) \ge \frac{1}{2\gamma} \|x^0 - x^*\|_2^2$

??

4. $f(x) = \frac{1}{2}\|Ax-b\|_2^2 \qquad g(x) = \|Ax-b\|_1$

a. $Hf(x) = A^TA \qquad 0 \leq Hf(x) \leq \lambda_{max}(A^TA)I$

So $f(x)$ is weakly convex and smooth w/ Lipschitz constant of $\nabla f \quad L = \lambda_{max}(A^TA)$

b. ~~Don't~~ final value $f =$ ~~0.00~~ 12. I do not believe it is optimal.

c. final value $obj = 3.7 \cdot 10^{-3}$

d. The convergence is not monotone b/c it goes back up a little around the 5th iteration.

e. $g(x) = \|Ax-b\|_1 = \sum_i |Ax_i - b|_1$

~~$g(x) \geq g(x) + g^T(y-x)$~~ $f(y) \geq f(x) + g^T(y-x)$

$$\partial|x|_1 = \begin{cases} 1 & x>0 \\ -1 & x<0 \\ [-1,1] & x=0 \end{cases}$$

So

$x>0$ or $x<0 \quad \nabla g(x) = \pm A$

$x=0 \qquad f(y) \geq f(x) + g^T(y-x)$

$$\frac{f(y)}{y} \geq g^T \longrightarrow \frac{|Ay-b|}{y} \geq g^T$$

$$A - b/y \geq g^T$$

$$\partial g(x) = \begin{cases} A^T & x>0 \\ -A^T & x<0 \\ [-A^T, A^T] & x=0 \end{cases}$$

$$= |A^T| \operatorname{Sgn}(A^Tg - b)$$

5. a. The final value of g is 0.423.

b. $g(x) = \frac{1}{n} \sum_{i=1}^{n} \log(1 + e^{-b_i a_i^T x})$

$$\nabla g(x) = \frac{1}{n} A \left(-b e^{-bA^T x}\right) / \left(1 + e^{-bA^T x}\right)$$

$$Hg(x) = \frac{1}{n} A \, \mathrm{diag}\left[\frac{b^2 \exp(-bA^T x)}{[1+\exp(-bA^T x)]^2}\right] A^T$$

$z^T Hg(x) z > 0 \rightarrow$ so $g(x)$ is convex

c. $\partial h(x) = \partial \lambda \|x\|_1 = \lambda \begin{cases} 1 & x > 0 \\ -1 & x < 0 \\ [-1, 1] & x = 0 \end{cases} = \lambda \, \mathrm{sgn}(x)$

$$\mathrm{prox}_{\gamma h}(y) = \mathrm{argmin}\left\{\frac{1}{2}\|x - y\|^2 + \gamma \lambda \|x\|_1\right\}$$

Let Dimension be 1D.

$x \geq 0$ $\quad \mathrm{argmin}\left\{\frac{1}{2}(x-y)^2 + \gamma\lambda x\right\}$

$= (x - y) + \gamma\lambda = 0$

$x = y - \gamma\lambda \qquad y > \gamma\lambda$

$x < 0$ $\quad \mathrm{argmin}\left\{\frac{1}{2}(x-y)^2 - \gamma\lambda x\right\}$

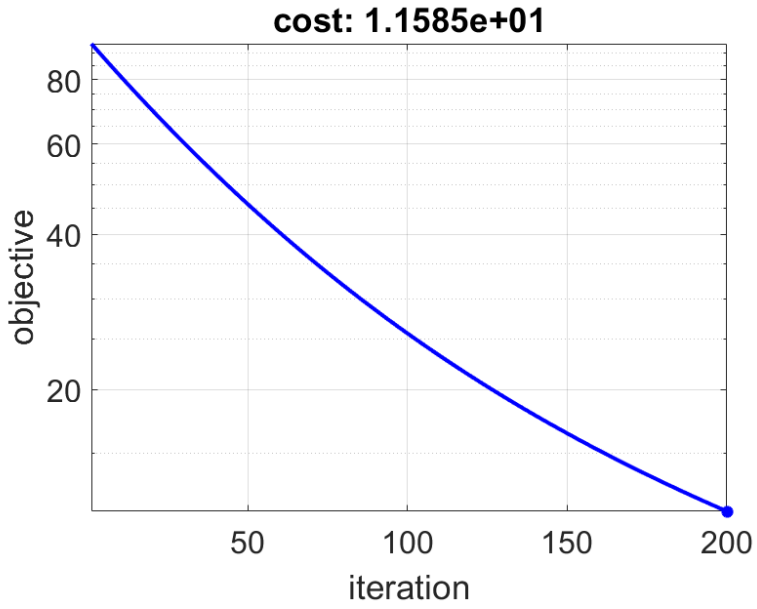
$= x - y - \gamma\lambda = 0 \rightarrow x = y + \gamma\lambda$

if $y < -\gamma\lambda$ then $x = y + \gamma\lambda$

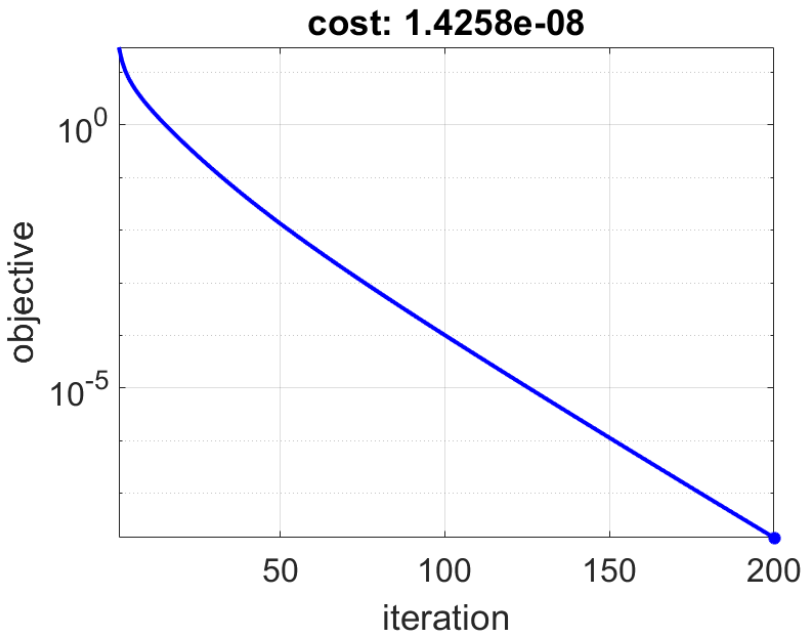
if $y > \gamma\lambda$ $\quad x = 0$.

$\mathrm{prox}_{\gamma h}(y) = |y - \gamma\lambda| \, \mathrm{sgn}(y)$

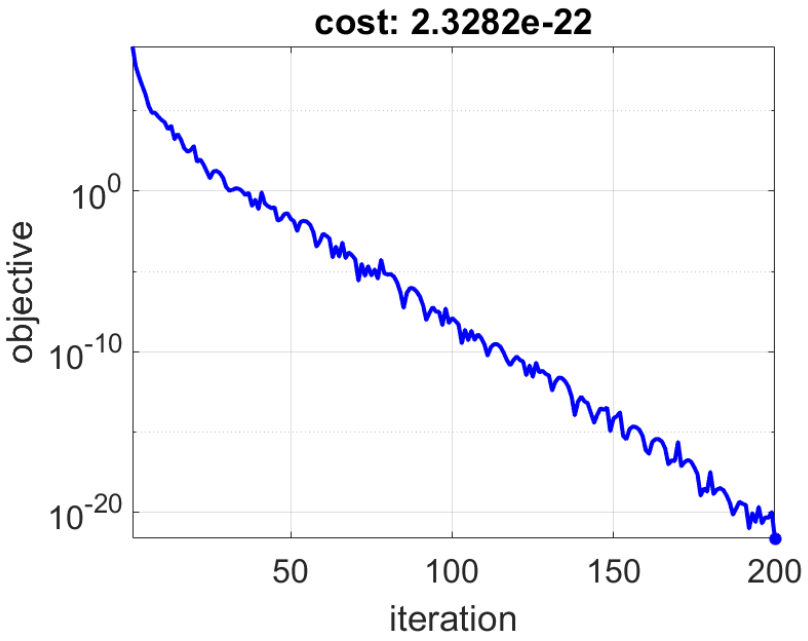
Problem 4b



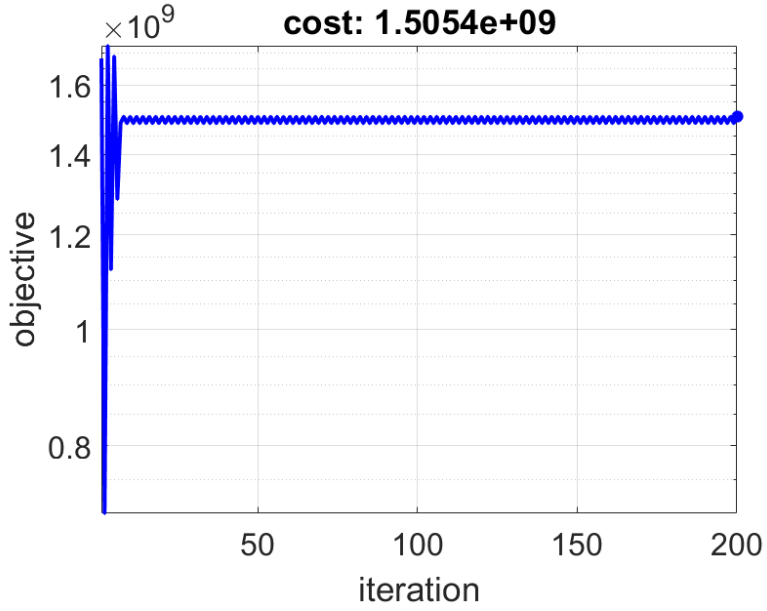
Problem 4c



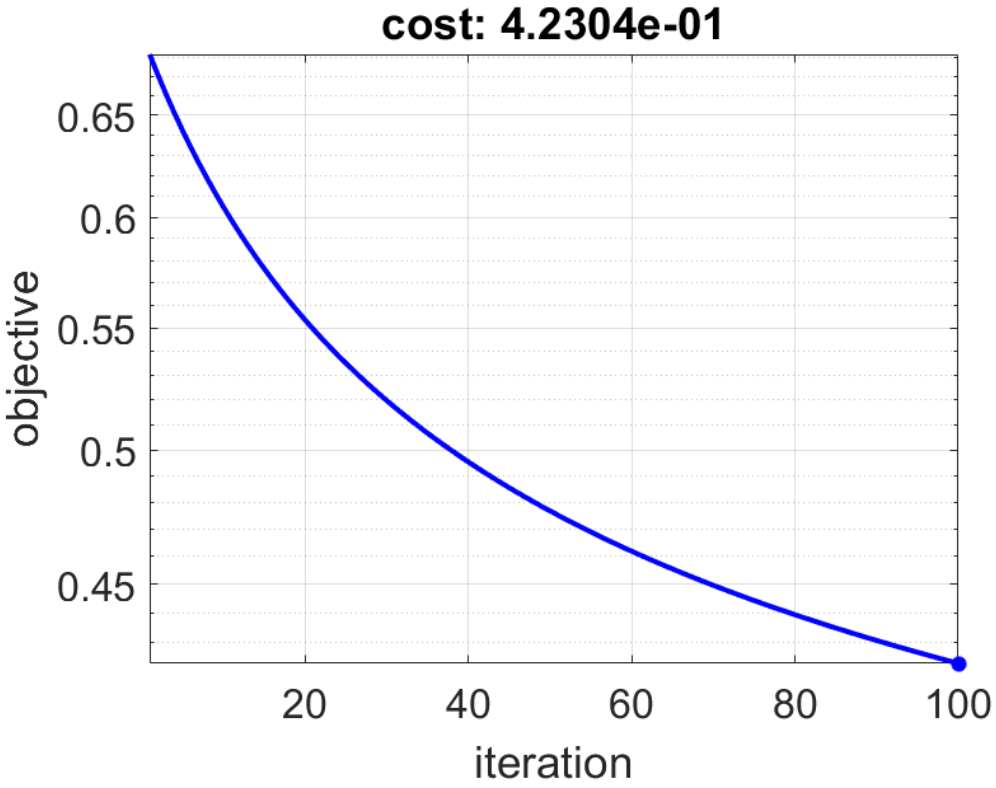
Problem 4d



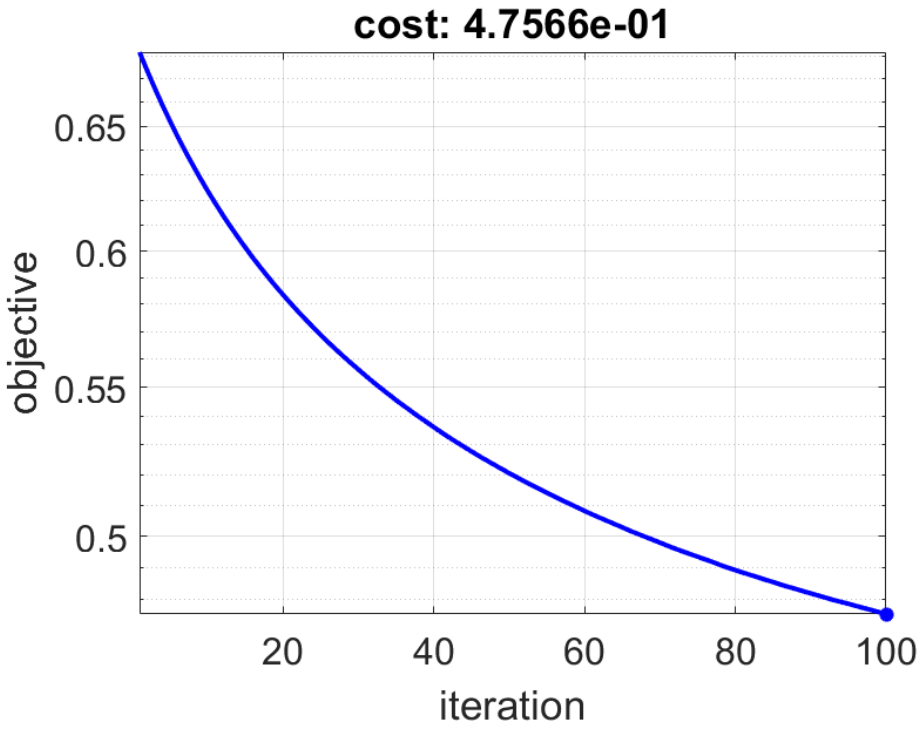
Problem 4e



Problem 5a

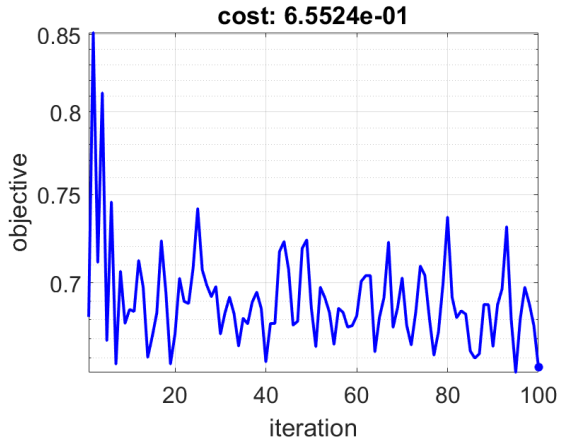


Problem 5d

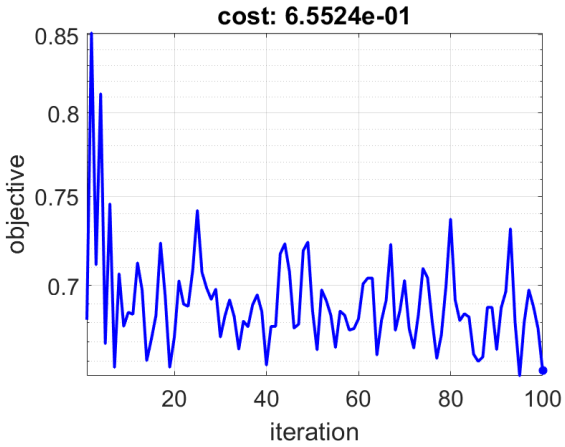


Problem 5e

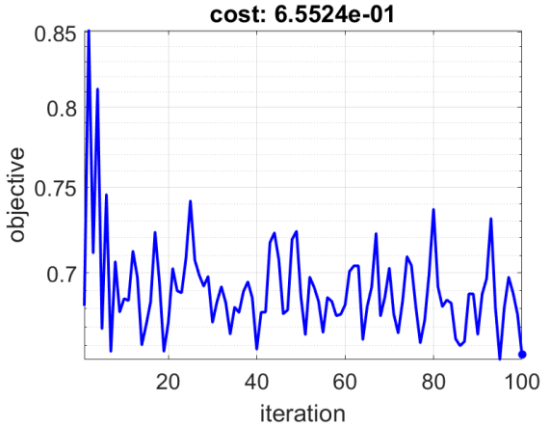
Mu = 0.1



Mu = 0.01



Mu=0.001



Contents


```
% GM with a fixed step
%
% Least squares: gradient method with fixed step
%
% U. S. Kamilov, CIG, WUSTL, 2021.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

prepare workspace

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear; close all; home;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

load the variables of the optimization problem

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

load('dataset.mat');

[m, n] = size(A); % m rows, n cols

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

set up the function and its gradient

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
stepSize = 1; % step-size of the gradient method
lambda = 0.02;

evaluateFunc = @(x) (1/2)*norm(A*x-b)^2;
evaluateGrad = @(x) A'*(A*x-b);

evaluate_g = @(x) lambda*norm(x,1);
evaluateGrad_g = @(y) abs(A').*sign(y); % This is the rule for subgrad of g
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

parameters of the gradient method

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
xInit = zeros(n, 1); % zero initialization
maxIter = 200; % maximum number of iterations
beta = 0.5; % Step reduction parameter
phi = 0.5;
thetaPast = 1;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

optimize

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% initialize
x = xInit;
xPast = x;

% keep track of cost function values
objVals = zeros(maxIter, 1);

% iterate
for iter = 1:maxIter


    % gradient at w
    grad = evaluateGrad(x);

%     % AGM
%     theta = (1+sqrt(1+4*thetaPast^2))/2;
%     beta_t = (thetaPast - 1)/theta;
%     s = x + beta_t*(x - xPast);
%     % Update x
%     xPast = x;
%
%     % update AGM
%     x = s - stepSize*evaluateGrad(s);
%
%     stepSize = 100;
%
%     % BLS
%     while (evaluateFunc(x - stepSize*evaluateGrad(x)) > evaluateFunc(x) ...
%                                     - phi*stepSize*norm(evaluateGrad(x))^2)
%     stepSize = beta*stepSize;
%
%     end

    % update GDM
    %xNext = x - stepSize*evaluateGrad(x);

    % SGD
    xNext = x - stepSize*evaluateGrad_g(x);

    % evaluate the objective
    funcNext = evaluateFunc(xNext);

    % store the objective and the classification error
    objVals(iter) = funcNext;

    fprintf('[%d/%d] [step: %.1e] [objective: %.1e] [norm(grad): %.1e]\n',...
        iter, maxIter, stepSize, objVals(iter), norm(grad));
```

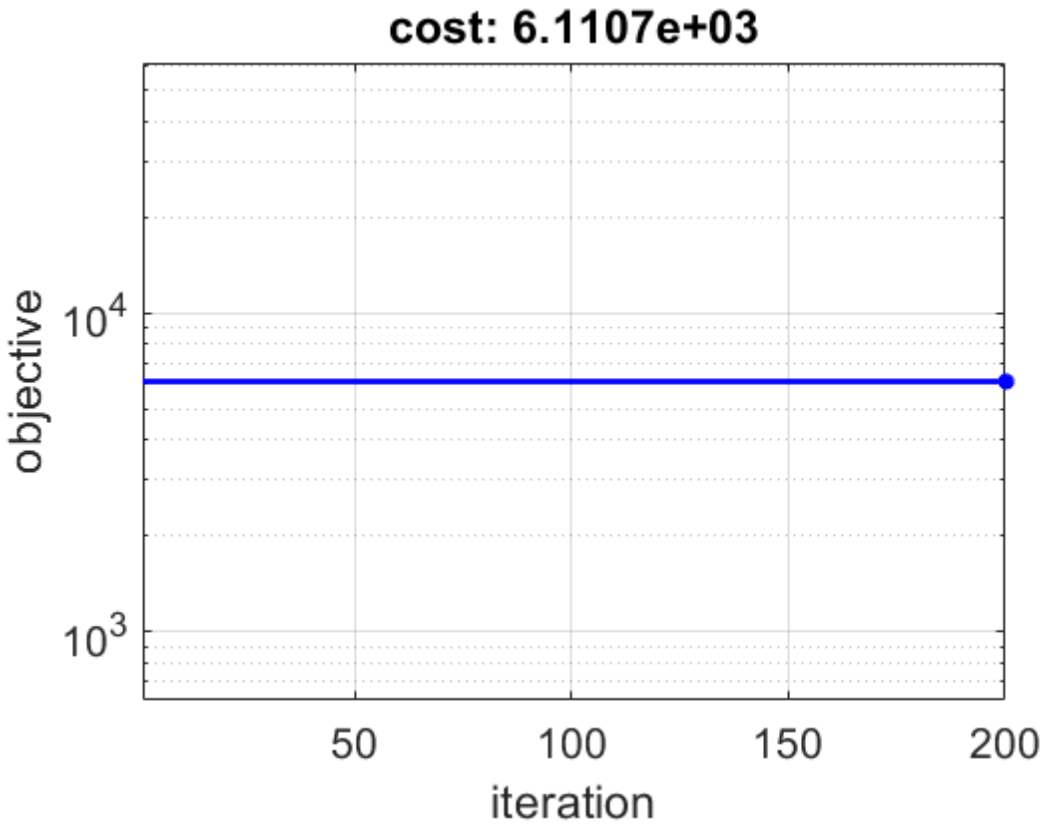
```
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % begin visualize data
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    % plot the evolution
    figure(1);
    set(gcf, 'Color', 'w');
    semilogy(1:iter, objVals(1:iter), 'b-',...
        iter, objVals(iter), 'b*', 'LineWidth', 2);
    grid on;
    axis tight;
    xlabel('iteration');
    ylabel('objective');
    title(sprintf('cost: %.4e', objVals(iter)));
    xlim([1 maxIter]);
    set(gca, 'FontSize', 16);
    drawnow;

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % end visualize data
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    % update w
    %thetaPast = theta;
    x = xNext;
end
```

```
[1/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.2e+02]
[2/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[3/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[4/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[5/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[6/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[7/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[8/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[9/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[10/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[11/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[12/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[13/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[14/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[15/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[16/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[17/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[18/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[19/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[20/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[21/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[22/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[23/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[24/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[25/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[26/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[27/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[28/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[29/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[30/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[31/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
```

```
[32/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[33/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[34/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[35/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[36/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[37/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[38/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[39/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[40/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[41/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[42/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[43/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[44/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[45/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[46/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[47/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[48/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[49/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[50/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[51/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[52/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[53/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[54/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[55/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[56/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[57/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[58/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[59/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[60/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[61/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[62/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[63/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[64/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[65/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[66/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[67/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[68/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[69/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[70/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[71/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[72/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[73/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[74/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[75/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[76/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[77/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[78/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[79/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[80/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[81/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[82/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[83/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[84/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[85/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[86/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[87/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[88/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[89/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[90/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[91/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[92/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
```

```
[93/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[94/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[95/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[96/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[97/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[98/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[99/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[100/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[101/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[102/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[103/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[104/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[105/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[106/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[107/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[108/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[109/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[110/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[111/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[112/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[113/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[114/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[115/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[116/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[117/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[118/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[119/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[120/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[121/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[122/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[123/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[124/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[125/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[126/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[127/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[128/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[129/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[130/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[131/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[132/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[133/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[134/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[135/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[136/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[137/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[138/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[139/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[140/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[141/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[142/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[143/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[144/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[145/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[146/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[147/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[148/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[149/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[150/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[151/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[152/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[153/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
```

```
[154/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[155/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[156/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[157/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[158/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[159/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[160/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[161/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[162/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[163/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[164/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[165/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[166/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[167/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[168/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[169/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[170/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[171/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[172/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[173/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[174/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[175/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[176/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[177/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[178/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[179/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[180/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[181/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[182/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[183/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[184/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[185/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[186/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[187/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[188/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[189/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[190/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[191/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[192/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[193/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[194/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[195/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[196/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[197/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[198/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[199/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
[200/200] [step: 1.0e+00] [objective: 6.1e+03] [norm(grad): 1.0e+03]
```



Published with MATLAB® R2020b

Contents


```
% Logistic Regression
%
% Gradient method with a fixed step
%
% U. S. Kamilov, CIG, WUSTL, 2021.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

prepare workspace

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear; close all; home;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

load the variables of the optimization problem

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

load('dataset.mat');

[p, n] = size(A);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

set up the function and its gradient

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
lambda = 0.02;
mu = 0.001;
evaluateFunc = @(x) (1/n)*sum(log(1+exp(-b.*(A'*x))));
evaluateGrad = @(x) (1/n)*A*(-b.*exp(-b.*(A'*x))./(1+exp(-b.*(A'*x))));

evaluateFunc_g = @(x) lambda*norm(x,1);
evaluateGrad_g = @(x) lambda*sign(x);
evaluateFunc_gsmooth = @(x) (x.^2/(2*mu))*(norm(x,1) < mu) ...
                          + (abs(x) - mu/2)*(norm(x,1) > mu);
evaluateGrad_gsmooth = @(x) (x/mu)*(norm(x,1) < mu) + (sign(x))*(norm(x,1) > mu);

evaluateFunc_h = @(x) evaluateFunc(x) + evaluateFunc_g(x);
evaluateGrad_h = @(x) evaluateGrad(x) + evaluateGrad_g(x);
evaluateGrad_hsmooth = @(x) evaluateGrad(x) + evaluateGrad_gsmooth(x);
```

```
prox_g = @(y) (abs(y) - stepSize*lambda).*sign(y);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

parameters of the gradient method

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

xInit = zeros(p, 1); % zero initialization
stepSize = 0.1; % step-size of the gradient method
tol = 1e-6; % stopping tolerance
maxIter = 100; % maximum number of iterations
thetaPast = 1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

optimize

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% initialize
x = xInit;
xPast = x;
% keep track of cost function values
objVals = zeros(maxIter, 1);

% iterate
for iter = 1:maxIter

    % gradient at w
    grad = evaluateGrad(x);

    % update
    %xNext = x - stepSize*grad;

    % SGD
    %xNext = x - stepSize*evaluateGrad_h(x);

    %AGM
    theta = (1+sqrt(1+4*thetaPast^2))/2;
    beta_t = (thetaPast - 1)/theta;
    s = x + beta_t*(x - xPast);
    xNext = prox_g(s - stepSize*evaluateGrad_hsmooth(s));


    % evaluate the objective
    funcNext = evaluateFunc(xNext);

    % store the objective and the classification error
    objVals(iter) = funcNext;

    fprintf('[%d/%d] [step: %.1e] [objective: %.1e] [norm(grad): %.1e]\n',...
        iter, maxIter, stepSize, objVals(iter), norm(grad));

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % begin visualize data
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    % plot the evolution
    figure(1);
```

```
    set(gcf, 'Color', 'w');
    semilogy(1:iter, objVals(1:iter), 'b-',...
        iter, objVals(iter), 'b*', 'LineWidth', 2);
    grid on;
    axis tight;
    xlabel('iteration');
    ylabel('objective');
    title(sprintf('cost: %.4e', objVals(iter)));
    xlim([1 maxIter]);
    set(gca, 'FontSize', 16);
    drawnow;

    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % end visualize data
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    % check stopping criterion
    if(norm(grad) < tol)
        break;
    end

    % update w
    thetaPast = theta;
    xPast = x;
    x = xNext;
end

% save for plotting
cost_gm = objVals;
save('plotfile.mat', 'cost_gm');
```

```
Unrecognized function or variable 'stepSize'.

Error in gm>@(y)(abs(y)-stepSize*lambda).*sign(y) (line 39)
prox_g = @(y) (abs(y) - stepSize*lambda).*sign(y);

Error in gm (line 75)
    xNext = prox_g(s - stepSize*evaluateGrad_hsmooth(s));
```

Published with MATLAB® R2020b