Topic: heavy ball method and Nesterov's accelerated method (Ch. 4)

Minimize a quadratic objective $f(x) = \frac{1}{2}x^T Ax$ with some first-order methods, generating the problems using the following MATLAB code fragment (or its equivalent in another language of your choice) to generate a Hessian with eigenvalues in the range $[m, L]$.

```
mu=0.01; L=1; kappa=L/mu;
n=100;
A = randn(n,n); [Q,R]=qr(A);
D=rand(n,1); D=10.^{D}; Dmin=min(D); Dmax=max(D);
D=(D-Dmin)/(Dmax-Dmin);
D = mu + D*(L-mu);
A = Q'*diag(D)*Q;
epsilon=1.e-6;
kmax=1000;
x0 = randn(n,1); % different x0 for each trial
```

Problem (Exercise 2 from Wright and Recht, Ch. 4) Run the code in each case until $f(x_k) \leq \epsilon$ for tolerance $\epsilon = 10^{-6}$. Implement the following methods.

- Steepest descent with $\alpha_k \equiv 2/(m + L)$
- Steepest descent with $\alpha_k \equiv 1/L$
- Steepest descent with exact line search
- Heavy-ball method, with $\alpha = 4/(\sqrt{L} + \sqrt{m})^2$ and $\beta = \left((\sqrt{L} - \sqrt{m})/(\sqrt{L} + \sqrt{m})\right)^2$
- Nesterov's optimal method, with $\alpha = 1/L$ and $\beta = (\sqrt{L} - \sqrt{m})/(\sqrt{L} + \sqrt{m})$

(a) Tabulate the average number of iterations required, over 10 random starts.

(b) Draw a plot of the convergence behavior on a typical run, plotting iteration number against $\log_{10}(f(x_k) - f(x_*))$. (Use the same figure, with different colors for each algorithm.)

(c) Discuss your results, noting in particular whether the worst-case convergence analysis is reflected in the practical results.