HOMEWORK 2. ACCELERATED GRADIENT DESCENT

Yutong He **Kun Yuan**

September 27, 2023

Attention: Turn in your homework at the beginning of our lecture on Oct. 10, 2023

1 Equivalence of iterations

(Exercise 2 in notes-chapter2) Prove that with proper initialization, the NAG iterations

$$\begin{aligned} y_{k-1} =& x_{k-1} + \beta(x_{k-1} - x_{k-2}), \\ x_k =& y_{k-1} - \gamma \nabla f(y_{k-1}), \end{aligned}$$

with parameter choices

$$\beta = \frac{\sqrt{L} - \sqrt{\mu}}{\sqrt{L} + \sqrt{\mu}}, \quad \gamma = \frac{1}{L},$$

are equivalent to the following iterations:

$$\begin{aligned} y_{k-1} =& \frac{\sqrt{L}}{\sqrt{L} + \sqrt{\mu}} x_{k-1} + \frac{\sqrt{\mu}}{\sqrt{L} + \sqrt{\mu}} v_{k-1}, \\ x_k =& y_{k-1} - \frac{1}{L} \nabla f(y_{k-1}), \\ v_k =& \left(1 - \frac{\sqrt{\mu}}{\sqrt{L}}\right) v_{k-1} + \frac{\sqrt{L}}{\sqrt{\mu}} x_k + \left(\frac{\sqrt{\mu}}{\sqrt{L}} - \frac{\sqrt{L}}{\sqrt{\mu}}\right) y_{k-1}. \end{aligned}$$

2 Deriving Anderson iterations

(Exercise 3 in notes-chapter2) Consider the following constrained convex optimization problem:

$$\min_{\alpha \in \mathbb{R}^m} \frac{1}{2}\|A\alpha\|_2^2, \quad \text{s.t. } \alpha^\top \mathbf{1} = 1,$$

where $A \in \mathbb{R}^{d\times m}$ has full column rank (which implies $A^\top A$ is non-singular). According to convex optimization theory, the optimal solution $\alpha^\star$ must satisfy the following KKT conditions:

$$\exists\ \lambda^\star \in \mathbb{R},\ \text{s.t.} \quad \begin{cases} \nabla_\alpha \mathcal{L}(\alpha^\star, \lambda^\star) = 0, \\ (\alpha^\star)^\top \mathbf{1} = 1, \end{cases}$$

where the Lagrange function $\mathcal{L}(\alpha, \lambda) := \frac{1}{2}\|A\alpha\|_2^2 + \lambda(\alpha^\top \mathbf{1} - 1)$. Prove that the optimal solution is

$$\alpha^\star = \frac{(A^\top A)^{-1}\mathbf{1}}{\mathbf{1}^\top (A^\top A)^{-1}\mathbf{1}}.$$