Modification of proof

September 1, 2022

In this note we fix a little typo involving the proof for the first-order potential reduction. Recall that in the note we claim the following relation

$$\mathbf{H} \succeq -\frac{\rho}{f(\mathbf{x})^2}\nabla f(\mathbf{x})\nabla f(\mathbf{x})^\top = -\frac{2\rho}{\|\mathbf{A}\mathbf{x}\|^2}\mathbf{A}^\top\mathbf{A}\mathbf{x}\mathbf{x}^\top\mathbf{A}^\top\mathbf{A} \succeq -2\rho\gamma$$

but since $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x}\|^2$, we actually have

$$-\frac{\rho}{f(\mathbf{x})^2}\nabla f(\mathbf{x})\nabla f(\mathbf{x})^\top = -\frac{4\rho}{\|\mathbf{A}\mathbf{x}\|^4}\mathbf{A}^\top\mathbf{A}\mathbf{x}\mathbf{x}^\top\mathbf{A}^\top\mathbf{A}$$

and the eigen-value bound becomes $\mathbf{x}$-dependent since

$$\begin{aligned}
& \max_{\mathbf{u}\neq 0}\left\langle \frac{\mathbf{u}}{\|\mathbf{u}\|^2}, \frac{4\rho}{\|\mathbf{A}\mathbf{x}\|^4}\mathbf{A}^\top\mathbf{A}\mathbf{x}\mathbf{x}^\top\mathbf{A}^\top\mathbf{A}\mathbf{u}\right\rangle \\
= \; & \max_{\mathbf{u}\neq 0}\frac{4\rho(\mathbf{x}\mathbf{A}^\top\mathbf{A}\mathbf{u})^2}{\|\mathbf{A}\mathbf{x}\|^4\|\mathbf{u}\|^2} \\
\leq \; & \frac{2\rho}{\frac{1}{2}\|\mathbf{A}\mathbf{x}\|^2}\max_{\mathbf{u}\neq 0}\frac{\|\mathbf{A}\mathbf{u}\|^2}{\|\mathbf{u}\|^2} \leq \frac{2\rho\gamma}{f(\mathbf{x})}.
\end{aligned}$$

By the modified bound and appealing to (3) in the original proof, we have

$$\frac{1}{2}\langle \mathbf{d}, \mathbf{H}\mathbf{d}\rangle \geq -\frac{\rho\gamma}{f(\mathbf{x})}\|\mathbf{d}\|^2 \geq -\frac{\rho\gamma\beta^2}{f(\mathbf{x})}.$$

Though the bound is worse, it does not matter since it is a high-order term of β.

Therefore,

$$\phi(\mathbf{x}+\mathbf{d}) - \phi(\mathbf{x}) \leq -\beta + \frac{\rho\gamma\beta^2}{2f(\mathbf{x})} + \frac{\beta^2}{2} + \frac{\rho\gamma\beta^2}{f(\mathbf{x})} \leq -\beta + \frac{3\rho\gamma\beta^2}{2f(\mathbf{x})} + \frac{\beta^2}{2}$$

and the rest of the results still apply to guarantee the $\mathcal{O}\big(\frac{1}{\varepsilon}\log(\frac{1}{\varepsilon})\big)$ result.