Notes on Strong Convexity and Smoothness

For any twice differentiable function $f : \mathbb{R}^d \mapsto \mathbb{R}$, by Taylor expansion and mean value theorem, we have

$$f(\mathbf{w}) = f(\mathbf{v}) + (\mathbf{w}-\mathbf{v})^T \nabla f(\mathbf{v}) + \frac{1}{2}(\mathbf{w}-\mathbf{v})^T \nabla^2 f(\mathbf{u})(\mathbf{w}-\mathbf{v}) \ ,$$

where $\mathbf{u} = a\mathbf{w} + (1-a)\mathbf{v}$ for some $a \in [0,1]$.

Strong Convexity. To establish strong convexity, it is sufficient to show: there is some $\alpha > 0$ such that for any $\mathbf{u}$,

$$\nabla^2 f(\mathbf{u}) - \alpha \mathbb{I} \succeq \mathbf{0} \ , \tag{1}$$

where $\mathbb{I}$ is the identity matrix, i.e., the matrix $\nabla^2 f(\mathbf{u}) - \alpha\mathbb{I}$ needs to be positive semi-definite.

Once we have shown this, we have shown strong convexity of f because we have

$$(\mathbf{w}-\mathbf{v})^T(\nabla^2 f(\mathbf{u}) - \alpha\mathbb{I})(\mathbf{w}-\mathbf{v}) \geq 0 \quad \Rightarrow \quad (\mathbf{w}-\mathbf{v})^T \nabla^2 f(\mathbf{u})(\mathbf{w}-\mathbf{v}) \geq \alpha \|\mathbf{w}-\mathbf{v}\|_2^2$$

so that

$$\begin{aligned} f(\mathbf{w}) &= f(\mathbf{v}) + (\mathbf{w}-\mathbf{v})^T \nabla f(\mathbf{v}) + \frac{1}{2}(\mathbf{w}-\mathbf{v})^T \nabla^2 f(\mathbf{u})(\mathbf{w}-\mathbf{v}) \\ &\geq f(\mathbf{v}) + (\mathbf{w}-\mathbf{v})^T \nabla f(\mathbf{v}) + \frac{\alpha}{2}\|\mathbf{w}-\mathbf{v}\|_2^2 \ , \end{aligned}$$

which is the definition of strong convexity.

Smoothness. To establish smoothness, it is sufficient to show: there is some $\beta < \infty$ such that for any $\mathbf{u}$,

$$\beta\mathbb{I} - \nabla^2 f(\mathbf{u}) \succeq \mathbf{0} \ , \tag{2}$$

where $\mathbb{I}$ is the identity matrix, i.e., the matrix $\beta\mathbb{I} - \nabla^2 f(\mathbf{u})$ needs to be positive semi-definite.

Once we have shown this, we have shown smoothness of f because we have

$$(\mathbf{w}-\mathbf{v})^T(\beta\mathbb{I} - \nabla^2 f(\mathbf{u}))(\mathbf{w}-\mathbf{v}) \geq 0 \quad \Rightarrow \quad (\mathbf{w}-\mathbf{v})^T \nabla^2 f(\mathbf{u})(\mathbf{w}-\mathbf{v}) \leq \beta \|\mathbf{w}-\mathbf{v}\|_2^2$$

so that

$$\begin{aligned} f(\mathbf{w}) &= f(\mathbf{v}) + (\mathbf{w}-\mathbf{v})^T \nabla f(\mathbf{v}) + \frac{1}{2}(\mathbf{w}-\mathbf{v})^T \nabla^2 f(\mathbf{u})(\mathbf{w}-\mathbf{v}) \\ &\leq f(\mathbf{v}) + (\mathbf{w}-\mathbf{v})^T \nabla f(\mathbf{v}) + \frac{\beta}{2}\|\mathbf{w}-\mathbf{v}\|_2^2 \ , \end{aligned}$$

which is the definition of smoothness.

Additional aspects:

(1) Recall that a $d \times d$ matrix is positive semi-definite if for any $z \in \mathbb{R}^d$, $z^T A z \geq 0$.

(2) When $d = 1$, $f''(u) \geq \alpha > 0$ for any u implies strong convexity of f.

(3) When $d = 1$, $f''(u) \leq \beta < \infty$ for any u implies smoothness of f.