Problem 1

Solution

Let $\gamma(s)$ be the line connecting u and v, i.e. $\gamma(s) = v + s(u - v)$. By integrating along $\gamma(s)$, we have

$$\begin{aligned} f(u) = f(v) + \int_0^1 \nabla^\intercal f(\gamma(s)) \frac{d\gamma(s)}{ds} ds &= f(v) + \int_0^1 \nabla^\intercal f(v + s(u-v))(u-v) ds \\ &= f(v) + \int_0^1 \left(\nabla^\intercal f(v + s(u-v)) + \nabla^\intercal f(v) - \nabla^\intercal f(v)\right)(u-v) ds \\ &= f(v) + \nabla^\intercal f(v)(u-v) + \int_0^1 \left(\nabla^\intercal f(v + s(u-v)) - \nabla^\intercal f(v)\right)(u-v) ds \\ &\leq f(v) + \nabla^\intercal f(v)(u-v) + \int_0^1 \left|\left|\nabla^\intercal f(v + s(u-v)) - \nabla^\intercal f(v)\right|\right| \left|\left|u-v\right|\right| ds \\ &\leq f(v) + \nabla^\intercal f(v)(u-v) + \int_0^1 sL \left|\left|u-v\right|\right| \cdot \left|\left|u-v\right|\right| ds \\ &= f(v) + (u-v)^\intercal \nabla f(v) + \frac{1}{2} L \left|\left|u-v\right|\right|^2. \end{aligned}$$

Problem 2

Solution

Starting from any point w in $\mathbb{R}^d$, we can follow the direction of the gradient of f and reach a local minima w^* such that $\nabla f(w^*) = 0$, i.e. we found a regular curve $\gamma : [0,1] \mapsto \mathbb{R}^n$ such that $\gamma(0) = w^*$, $\gamma(1) = w$, and the velocity vector at $\gamma(s)$ is $\frac{d\gamma(s)}{ds} = c(s)\nabla f(\gamma(s))$, where $c(s) > 0$ is a scalar. We denote $\gamma_s(s) = \frac{d\gamma(s)}{ds}$ the velocity. Then, we integrate along $\gamma(s)$. Since, $f(w^*) > 0$, we have

$$2L \cdot f(w) = 2L \cdot \left(f(w^*) + \int_0^1 \nabla^\intercal f(\gamma(\mathbf{s}))\gamma_s(\mathbf{s}) d\mathbf{s} \right) \geq 2L \cdot \int_0^1 \nabla^\intercal f(\gamma(\mathbf{s}))\gamma_s(\mathbf{s}) d\mathbf{s}.$$

Since $\nabla f(w)$ is L-Lipschitz continuous, and $\nabla^\intercal f(\gamma(\mathbf{s})) \cdot \gamma_s(\mathbf{s}) = c(\mathbf{s}) \nabla^\intercal f(\gamma(\mathbf{s})) \cdot \nabla f(\gamma(\mathbf{s}))$ is always non-negative, we have that the increament of $||\nabla f(w)||^2$ is bounded, and

$$||\nabla f(w)||^2 = ||\nabla f(w)||^2 - ||\nabla f(w^*)||^2 \leq \int_0^1 2L \cdot \nabla^\intercal f(\gamma(\mathbf{s}))\gamma_s(\mathbf{s}) d\mathbf{s}.$$

Thus, we have

$$||\nabla f(w)||^2 \leq 2L \cdot f(w).$$