Notes for Lecture 13

Scribe: Tingkai Jia

1 Properties of Minimax Optimization

Property 1. *In minimax optimization, we have*

$$\max_{\mathbf{y}\in\mathcal{Y}}\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x},\mathbf{y}) \le \min_{\mathbf{x}\in\mathcal{X}}\max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{x},\mathbf{y})$$

Proof. Let

$$\begin{aligned}\mathbf{y}^* &= \arg\max_{\mathbf{y}\in\mathcal{Y}}\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x},\mathbf{y})\\ \mathbf{x}^* &= \arg\min_{\mathbf{x}\in\mathcal{X}}\max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{x},\mathbf{y}),\end{aligned}$$

we naturally have

$$\begin{aligned}\max_{\mathbf{y}\in\mathcal{Y}}\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x},\mathbf{y}) &= \min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x},\mathbf{y}^*) \le f(\mathbf{x}^*,\mathbf{y}^*)\\ \min_{\mathbf{x}\in\mathcal{X}}\max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{x},\mathbf{y}) &= \max_{\mathbf{y}\in\mathcal{Y}} f(\mathbf{x}^*,\mathbf{y}) \ge f(\mathbf{x}^*,\mathbf{y}^*),\end{aligned}$$

thus we finish the proof. □

2 Convergence Rates of GDA

Theorem 1. *If f is l-smooth and convex-concave, let the diameter of $\mathcal{X}$ and $\mathcal{Y}$ be R. Then for fixed t with $\eta = \frac{R}{L\sqrt{t}}$, where $L = 2lR + \|\nabla f(\mathbf{x}_0,\mathbf{y}_0\|_2$, we have*

$$\max_{\mathbf{y}\in\mathcal{Y}} f\left(\frac{1}{t}\sum_{k=1}^{t}\mathbf{x}_k,\mathbf{y}\right) - \min_{\mathbf{x}\in\mathcal{X}} f\left(\mathbf{x},\frac{1}{t}\sum_{k=1}^{t}\mathbf{y}_k\right) \le \frac{2LR}{\sqrt{t}}.$$

Proof. It follows that

$$\begin{aligned}f(\mathbf{x}_k,\mathbf{y}) - f(\mathbf{x},\mathbf{y}_k) &= f(\mathbf{x}_k,\mathbf{y}) - f(\mathbf{x}_k,\mathbf{y}_k) + f(\mathbf{x}_k,\mathbf{y}_k) - f(\mathbf{x},\mathbf{y}_k)\\ &\le \nabla_{\mathbf{y}} f(\mathbf{x}_k,\mathbf{y}_k)^\top(\mathbf{y}-\mathbf{y}_k) + \nabla_{\mathbf{x}} f(\mathbf{x}_k,\mathbf{y}_k)^\top(\mathbf{x}_k-\mathbf{x})\\ &= \frac{1}{\eta}(\tilde{\mathbf{y}}_{k+1}-\mathbf{y}_k)^\top(\mathbf{y}-\mathbf{y}_k) + \frac{1}{\eta}(\mathbf{x}_k-\tilde{\mathbf{x}}_{k+1})^\top(\mathbf{x}_k-\mathbf{x})\\ &= \frac{1}{2\eta}\left[\|\tilde{\mathbf{y}}_{k+1}-\mathbf{y}_k\|_2^2 + \|\mathbf{y}_k-\mathbf{y}\|_2^2 - \|\tilde{\mathbf{y}}_{k+1}-\mathbf{y}\|_2^2\right] +\\ &\qquad \frac{1}{2\eta}\left[\|\mathbf{x}_k-\tilde{\mathbf{x}}_{k+1}\|_2^2 + \|\mathbf{x}-\mathbf{x}_k\|_2^2 - \|\mathbf{x}-\tilde{\mathbf{x}}_{k+1}\|_2^2\right]\\ &\le \frac{1}{2\eta}\left[\eta^2L^2 + \|\mathbf{y}_k-\mathbf{y}\|_2^2 - \|\tilde{\mathbf{y}}_{k+1}-\mathbf{y}\|_2^2\right] +\\ &\qquad \frac{1}{2\eta}\left[\eta^2L^2 + \|\mathbf{x}-\mathbf{x}_k\|_2^2 - \|\mathbf{x}-\tilde{\mathbf{x}}_{k+1}\|_2^2\right].\end{aligned}$$

With the property of projection

$$\|\mathbf{x} - \mathcal{P}_{\mathcal{C}}(\mathbf{x})\|_2^2 + \|\mathbf{z} - \mathcal{P}_{\mathcal{C}}(\mathbf{x})\|_2^2 \leq \|\mathbf{x} - \mathbf{z}\|_2^2, \quad \mathbf{z} \in \mathcal{C}$$

we have

$$\begin{aligned} \|\tilde{\mathbf{y}}_{k+1} - \mathbf{y}_{k+1}\|_2^2 + \|\mathbf{y} - \mathbf{y}_{k+1}\|_2^2 &\leq \|\tilde{\mathbf{y}}_{k+1} - \mathbf{y}\|_2^2 \\ \|\mathbf{y} - \mathbf{y}_{k+1}\|_2^2 &\leq \|\tilde{\mathbf{y}}_{k+1} - \mathbf{y}\|_2^2, \end{aligned}$$

then we obtain

$$\begin{aligned} f(\mathbf{x}_k, \mathbf{y}) - f(\mathbf{x}, \mathbf{y}_k) &\leq \frac{1}{2\eta}\left[\eta^2 L^2 + \|\mathbf{y}_k - \mathbf{y}\|_2^2 - \|\tilde{\mathbf{y}}_{k+1} - \mathbf{y}\|_2^2\right] + \\ &\quad \frac{1}{2\eta}\left[\eta^2 L^2 + \|\mathbf{x} - \mathbf{x}_k\|_2^2 - \|\mathbf{x} - \tilde{\mathbf{x}}_{k+1}\|_2^2\right] \\ &\leq \frac{1}{2\eta}\left[\eta^2 L^2 + \|\mathbf{y}_k - \mathbf{y}\|_2^2 - \|\mathbf{y}_{k+1} - \mathbf{y}\|_2^2\right] + \\ &\quad \frac{1}{2\eta}\left[\eta^2 L^2 + \|\mathbf{x} - \mathbf{x}_k\|_2^2 - \|\mathbf{x} - \mathbf{x}_{k+1}\|_2^2\right], \end{aligned}$$

after apply it recursively, we obtain

$$\begin{aligned} \frac{1}{t}\sum_{k=1}^{t}\left(f(\mathbf{x}_k, \mathbf{y}) - f(\mathbf{x}, \mathbf{y}_k)\right) &\leq \frac{1}{2\eta}\left[\eta^2 L^2 + \frac{\|\mathbf{y}_1 - \mathbf{y}\|_2^2 - \|\mathbf{y}_{t+1} - \mathbf{y}\|_2^2}{t}\right] + \\ &\quad \frac{1}{2\eta}\left[\eta^2 L^2 + \frac{\|\mathbf{x} - \mathbf{x}_1\|_2^2 - \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2}{t}\right] \\ &\leq \frac{1}{2\eta}\left[\eta^2 L^2 + \frac{\|\mathbf{y}_1 - \mathbf{y}\|_2^2}{t}\right] + \frac{1}{2\eta}\left[\eta^2 L^2 + \frac{\|\mathbf{x} - \mathbf{x}_1\|_2^2}{t}\right] \\ &\leq \frac{1}{2\eta}\left(\eta^2 L^2 + \frac{R^2}{t}\right) + \frac{1}{2\eta}\left(\eta^2 L^2 + \frac{R^2}{t}\right) \\ &\leq \eta L^2 + \frac{R^2}{t\eta} \\ &\leq \frac{2LR}{\sqrt{t}}. \end{aligned}$$

□