[CS3958: Lecture 9] Mirror Descent

Instructor: Chihao Zhang, Scribed by Yulin Wang

December 7, 2022

1 Quadratic Regularizer

Recall that the updating rule of the gradient descent algorithm is $x_{t+1} = \Pi_V(x_t - \eta\nabla f(x_t))$. We have the following result:

Proposition 1 *The step $x_{t+1} = \Pi_V(x_t - \eta\nabla f(x_t))$ is equivalent to*

$$x_{t+1} = \arg\min_{x^* \in V} \langle \nabla f(x_t), x^* - x_t \rangle + \frac{1}{2\eta}\|x^* - x_t\|^2.$$

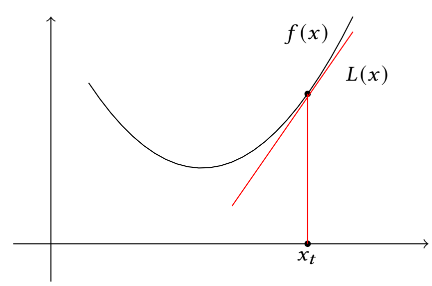
Proof.

$$\begin{aligned}\Pi_V(x_t - \eta\nabla f(x_t)) &= \arg\min_{x^* \in V} \|x_t - \eta\nabla f(x_t) - x^*\|^2 \\ &= \arg\min_{x^* \in V} \langle x_t - \eta\nabla f(x_t) - x^*, x_t - \eta\nabla f(x_t) - x^* \rangle \\ &= \arg\min_{x^* \in V} \langle x^*, x^* \rangle - 2\langle x^*, x_t - \eta\nabla f(x_t) \rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}\arg\min_{x^* \in V} \langle \nabla f(x_t), x^* - x_t \rangle + \frac{1}{2\eta}\|x^* - x_t\|^2 &= \arg\min_{x^* \in V} 2\eta\langle \nabla f(x_t), x^* \rangle + \langle x^*, x^* \rangle - 2\langle x^*, x_t \rangle \\ &= \arg\min_{x^* \in V} \langle x^*, x^* \rangle - 2\langle x^*, x_t - \eta\nabla f(x_t) \rangle \\ &= \Pi_V(x_t - \eta\nabla f(x_t)).\end{aligned}$$

□



The intuition behind the updating rule is as follows: Since the only information we know about $f(x)$ at t-th step is the gradient $\nabla f(x_t)$. We locally use the linear function $L(x) \triangleq \langle \nabla f(x_t), x - x_t \rangle$ to approximate $f(x)$, and therefore we try to use $\arg\min_x L(x)$ to approximate $\arg\min_x f(x)$.

However, the approximation is not correct as $\min L(x) = -\infty$ as long as $\nabla f(x_t) \neq 0$ since $L(x) \approx f(x)$ only holds for $x \approx x_t$. Therefore, we add a *regularization function* $\frac{1}{2\eta}\|x - x_t\|^2$ to force that x is close to x_t.

The choice of the quadratic function here is, in some sense, mainly for simplicity and not necessarily optimal. Intuitively, if the our local approximator for $f(x)$ is more close to f, our algorithm is more efficient. As a result, we can choose some better regularizer once some information about f is known.

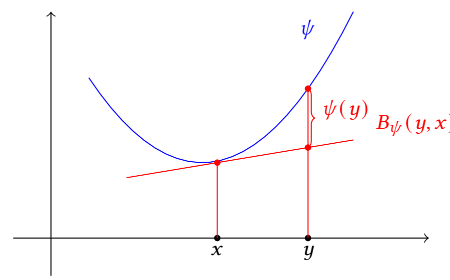
2 Mirror Descent

First, we fix some notations.

- $\psi : \mathbb{R}^n \to \mathbb{R}$ is a convex function.
- For any $x, y \in \mathbb{R}^n$, the *Bregman divergence* of ψ is defined as

$$B_\psi(y, x) = \psi(y) - [\psi(x) + \langle \nabla\psi(x), y - x \rangle] .$$

Intuitive, the Bregman divergence $B_\psi(y, x)$ is the difference between $\psi(y)$ and the linear approximation of $\psi(y)$ at $(x, \psi(x))$ (See the figure on the right).



For example, let $\psi(x) = \frac{1}{2}\|x\|^2$. Then

$$B_\psi(y, x) = \frac{1}{2}\|y\|^2 - \left[\frac{1}{2}\|x\|^2 + \langle x, y - x\rangle\right] = \frac{1}{2}\|x - y\|^2,$$

which is exactly the Euclidean distance between x, y.

Another important example is the *negative entropy* function. For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let $\psi(x) = \sum_{i=1}^n x_i \log x_i$. It is obvious that ψ is a convex function since

$$\nabla^2\psi(x)_{i,j} = \begin{cases} \frac{1}{x_i} & i = j; \\ 0 & o.w. \end{cases} \tag{1}$$

For a distribution $p \in \Delta_{n-1}$, the *entropy* of p is defined as

$$H(p) \triangleq -\sum_{i=1}^n p_i \log p_i.$$

Then the Bregman divergence of ψ is

$$\begin{aligned} B_\psi(y, x) &= \sum_{i=1}^n y_i \log y_i - \left(\sum_{i=1}^n (y_i - x_i)(1 + \log x_i)\right) \\ &= \sum_{i=1}^n y_i \log y_i - \sum_{i=1}^n y_i \log x_i - \sum_{i=1}^n y_i + \sum_{i=1}^n x_i. \end{aligned}$$

If we restrict x, y to Δ_{n-1}, then we have

$$\begin{aligned} B_\psi(y, x) &= \sum_{i=1}^n y_i \log \frac{y_i}{x_i} \\ &= D_{KL}(\mu_y \parallel \mu_x). \end{aligned}$$

where μ_x, μ_y are the distributions on $[n]$ such that $\mu_x(i) = x_i$, and $\mu_y(i) = y_i$.

Let μ, ν be distributions on the same probability space $\mathcal{X}$. Then the *Kullback-Leibler divergence*, i.e. *relative entropy* from μ to ν is defined as $D_{KL}(\nu \parallel \mu) = \sum_{x \in \mathcal{X}} \nu(x) \log \frac{\nu(x)}{\mu(x)}$. A simple interpretation of the KL divergence of ν from μ is the expected excess information content from using μ as a model when the actual distribution is ν.

2.1 *Updating Rule of Mirror Descent*

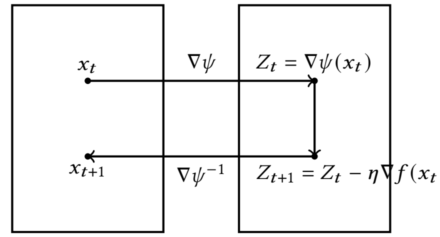
Now we generalize the quadratic regularizer in the gradient descent algorithm to Bregman divergences. Let $V = \mathbb{R}^n$. Fix a convex function ψ. Then the updating rule of the *mirror descent* algorithm is

$$x_{t+1} = \arg\min_{x \in V} \langle \nabla f(x_t), x - x_t \rangle + \frac{1}{\eta} B_\psi(x, x_t). \tag{2}$$

And the following proposition holds.

Proposition 2 *The updating rule 2 is equivalent to*

$$\nabla\psi(x_{t+1}) = \nabla\psi(x_t) - \eta\nabla f(x_t).$$



Proof. If ψ is chosen to be $\frac{1}{2}\|x\|^2$, then the equivalence holds naturally since $\nabla\psi$ is the identity function. We can prove the result when ψ is an arbitrary convex function.

Let $G(x) \triangleq \langle \nabla f(x_t), x - x_t \rangle + \frac{1}{\eta} B_\psi(x, x_t)$. We have

$$\begin{aligned}
\nabla G(x) = 0 &\Longleftrightarrow \nabla f(x_t) + \frac{1}{\eta}\nabla[\psi(x) - \psi(x_t) - \langle x - x_t, \nabla\psi(x_t)\rangle] \\
&\Longleftrightarrow \nabla f(x_t) + \frac{1}{\eta}(\nabla\psi(x) - \nabla\psi(x_t)) \\
&\Longleftrightarrow \nabla\psi(x) = \nabla\psi(x_t) - \nabla f(x_t).
\end{aligned}$$

□

Projection

Now we consider the constrained mirror descent, that is, the domain V is some closed convex subset of $\mathbb{R}^n$. For an arbitrary V, we need to restrict x_{t+1} into V in the updating step. It is worth noting that the projection here is with respect to Bregman divergence instead of Euclidean distance. So the updating rule 2 is equivalent to

- $Z_t = \nabla\psi(x_t)$;
- $Z_{t+1} = Z_t - \eta\nabla f(x_t)$;
- $x'_{t+1} = \nabla\psi^{-1}(Z_t)$;
- $x_{t+1} = \Pi_V^\psi(x'_{t+1}) \triangleq \arg\min_{x\in V} B_\psi(x, x'_{t+1})$.

The proof is omitted here since it is quite similar to that of Proposition 1.

2.2 *Applications of (Online) Mirror Descent*

Online learning

Applying the mirror descent algorithm to online learning settings in the previous lecture introduces the regret bound that

$$\begin{aligned}
R(T) &= \sum_{t=0}^{T-1} \ell_t(x_t) - \ell_t(x^*) \\
&\leq \sum_{t=0}^{T-1} \left[\frac{\phi(x_t) - \phi(x_{t+1})}{\eta_t} + \eta_t \cdot \sup_{\xi_t \in [x_t, x'_{t+1}]} \|\nabla f(x_t)\|^2_{\nabla^2\psi(\xi_t)^{-1}} \right],
\end{aligned}$$

where $\phi(x) = B_\psi(x^*, x)$. We take $\eta_t = \eta$ for all t. Then the regret is

For any matrix M, the *matrix norm* $\|\cdot\|_M$ is defined as $\|x\|_M \triangleq \sqrt{x^T M x}$.

$$R(T) \le \frac{\phi(x_0)}{\eta} + \eta \cdot \sum_{t=0}^{T-1} \sup_{\xi_t \in [x_t, x'_{t+1}]} \|\nabla f(x_t)\|^2_{\nabla^2 \psi(\xi_t)^{-1}}.$$

If $\psi = \frac{1}{2}\|x\|^2$, then the matrix bound is exactly 2-norm since $\nabla^2 \psi = I$, which implies exactly the regret bound in previous lectures.

Learning with expert advice

Applying the online mirror descent algorithm to the learning with expert advice problem, we obtain the following regret bound:

$$R(T) \le \frac{B_\psi(x^*, x_0)}{\eta} + \eta \sum_{t=0}^{T-1} \sup_{\xi_t \in [x_t, x'_{t+1}]} \|\ell_t\|^2_{\nabla^2 \psi(\xi_t)^{-1}}. \tag{3}$$

Recall that in this setting, $x_t \in \Delta_{n-1}$ for any t. If we take $\psi(x) = \sum_i x_i \log x_i$, then the diameter of Δ_{n-1} w.r.t ψ is

$$\text{diam}^\psi(\Delta_{n-1}) = O(\log n),$$

which means that the first term on RHS of Eq. 3 is no more that $\frac{O(\log n)}{\eta}$.

As for the second term, by Eq. 1, we have

$$\begin{aligned}
\eta \sum_{t=0}^{T-1} \sup_{\xi_t \in [x_t, x'_{t+1}]} \|\ell_t\|^2_{\nabla^2 \psi(\xi_t)^{-1}} &= \eta \sum_{t=0}^{T-1} \sup_{\xi_t \in [x_t, x'_{t+1}]} \sum_{i=1}^{n} \ell_t(i)^2 \xi_t(i) \\
&\le \eta \sum_{t=0}^{T-1} \sup_{\xi_t \in [x_t, x'_{t+1}]} \sum_{i=1}^{n} \xi_t(i) \\
&\le \eta \sum_{t=0}^{T-1} \sum_{i=1}^{n} x_t(i) \\
&= \eta T,
\end{aligned}$$

where we applied $\xi_t(i) \le x_t(i)$ since $\exists z \in [Z_{t+1}, Z_t]$, $\nabla \psi^{-1}(z) = \xi_t$ and $\nabla \psi$ is monotone increasing. Finally, we obtain the regret bound that

$$R(T) \le \frac{O(\log n)}{\eta} + \eta T \le O(\sqrt{\log n T})$$

by choosing $\eta = \sqrt{\frac{\log n}{T}}$.

The regret bound $\frac{1}{\eta} + \eta n T \le \sqrt{nT}$ in previous lectures is obtained by choosing $\psi = \frac{1}{2}\|x\|^2$. By changing the potential function ψ, we can balance between these two terms and obtain an optimal bound.

Multiplicative weights update

In this part, we focus on the combinatorial meaning of the mirror descent algorithm when choosing $\psi(x) = \sum_{i=1}^n x_i \log x_i$. In this case, $\nabla \psi(x) = (1 + \log x(1), 1 + \log x(2), \ldots, 1 + \log x(n))$. Plugging in the definition of ψ, the updating steps turn to

- $Z_t(i) = 1 + \log x_t(i)$;
- $Z_{t+1}(i) = 1 + \log x_t(i) - \eta \nabla \ell_t(i)$;

- $x'_{t+1} = x_t(i)e^{-\eta\ell_t(i)}$;
- $x_{t+1} = \Pi^{\psi}_{\Delta_{n-1}}(x'_{t+1}) = \arg\min_{x\in\Delta_{n-1}} \sum_{i=1}^n x(i)\log\frac{x(i)}{x'_{t+1}(i)}$. This constrained convex optimization problem can be solved by the method of Lagrange multiplier. It is not hard to obtain that

$$x_{t+1}(i) = \frac{x'_{t+1}(i)}{\sum_j x'_{t+1}(j)} = \frac{x_t(i)e^{-\eta\ell_t(i)}}{\sum_j x_t(j)e^{-\eta\ell_t(j)}}.$$

The above algorithm is called the *multiplicative weights update method*, which was discovered repeatedly in very diverse fields such as machine learning, optimization, theoretical computer science, and game theory.

Online Stochastic Mirror Descent (OSMD)

Let us return to the multi-armed bandit problem. We still fix $\psi(x) = \sum_{i=1}^n x_i \log x_i$ here. The regret bound is

$$\begin{aligned} R(T) &\le \frac{B_\psi(x^*, x_0)}{\eta} + \eta\mathbf{E}\left[\sum_{t=0}^{T-1} \sup_{\xi_t\in[x_t,x'_{t+1}]} \|\ell_t\|^2_{\nabla^2\psi(\xi_t)^{-1}}\right] && (4)\\ &\le \frac{\log n}{\eta} + \eta\mathbf{E}\left[\sum_{i=1}^n \left(\frac{1[A_t = i]\ell_t(i)}{x_t(i)}\right)^2 \ell_t(i)\right] && (5)\\ &\le \frac{\log n}{\eta} + \eta\mathbf{E}\left[\sum_{i=1}^n \frac{1}{x_t(i)} x_t(i)\right] && (6)\\ &= \frac{\log n}{\eta} + \eta T n && (7)\\ &\le \sqrt{n\log nT} && (8) \end{aligned}$$

by choosing $\eta = \sqrt{\frac{\log n}{nT}}$.