Fancy Bregman Game

Manfred Warmuth

University of California - Santa Cruz

December 13, 2011
Basics

- $F$ is convex function, $G$ its dual

$$G(f(x)) = x \cdot f(x) - F(x)$$

- $$\Delta_F(w, x) = F(w) - F(x) - (w - x) \cdot f(x) = F(w) + G(f(x)) - w \cdot f(x)$$
The game

The following minimax problem can be solved by a binary search on a concave maximization problem in $\alpha$:

$$
\max_{\alpha} \min_w \left( \sum_t \alpha_t \Delta_F(w, x_t) - \text{conv}(\alpha) \right)
$$

The $x_t$ are points. $w$ should be “close” to the $x_t$ as measured by a Bregman divergence $\Delta_F(w, x_t)$. The divergence is weighted by the coefficient $\alpha_t$. Finally we subtract a convex function in the coefficient vector $\alpha$

First optimize $w$:

$$
\frac{\partial \text{objective}}{\partial w} = \sum_t \alpha_t f(w) - \sum_t \alpha_t f(x_t) = 0
$$

$$
w^* = f^{-1}\left( \frac{\sum_t \alpha_t f(x_t)}{\sum_t \alpha_t} \right)
$$
Plugging in \( w = w^* \) and rewriting the objective

\[
\sum_t \alpha_t \Delta_F(w^*, x_t) - \text{conv}(\alpha)
\]

\[
= \sum_t \alpha_t F(w^*) + \sum_t \alpha_t G(f(x_t)) - w^* \cdot \sum_t \alpha_t f(x_t) - \text{conv}(\alpha)
\]

\[
= \left( \sum_t \alpha_t \right) \left( F(w^*) - w^* \cdot \frac{\sum_t \alpha_t f(x_t)}{\sum_t \alpha_t} \right) + \sum_t \alpha_t G(f(x_t)) - \text{conv}(\alpha)
\]

\[
= -\left( \sum_t \alpha_t \right) G(f(w^*)) + \sum_t \alpha_t G(f(x_t)) - \text{conv}(\alpha)
\]

\[
= \left( \sum_t \alpha_t \right) \left( \frac{\sum_t \alpha_t G(f(x_t))}{\sum_t \alpha_t} - G \left( \frac{\sum_t \alpha_t f(x_t)}{\sum_t \alpha_t} \right) \right) - \text{conv}(\alpha)
\]
Conclusion

\[ V = \max_{\alpha} \min_{w} \left( \sum_{t} \alpha_t \Delta F(w, x_t) - \text{conv}(\alpha) \right) \]

\[ = \max_{s} \max_{\alpha} V(s, \alpha), \]

where \( V(s, \alpha) = s \left( \frac{\sum_{t} \alpha_t G(f(x_t))}{s} - G \left( \frac{\sum_{t} \alpha_t f(x_t)}{s} \right) \right) - \text{conv}(\alpha) \)

- \( V(s, \alpha) \) is concave in \( \alpha \) and therefore can be maxed out
- Claim: \( V(s) := \max_{\alpha} V(s, \alpha) \) is not necessarily concave in \( s \), but probably has only two maxima
- If this is so, then the value of the entire game can be solved via binary searches on \( s \) while solving a concave problem in each iteration