May 10, 2021

$$f(x) = \frac{1}{n}\sum_{i=1}^{n} f_i(x)$$

for $k = 1, 2, \ldots$

$J \sim \text{Uniform}(\{1, \ldots, n\})$

$x_{k+1} = x_k - \alpha_k \nabla f_J(x_k)$

$$\mathbb{E}(\nabla f_J(x_k)) = \nabla f(x_k)$$

$$x_{k+1} = x_k - \alpha p_k \qquad p_k = \nabla f(x_k) + u_k$$

$$f(x) = \frac{1}{2}x^T A x + x^T b + c \Rightarrow \nabla f(x) = Ax + b$$

$$x_{k+1} - x_* = x_k - x_* - \alpha(Ax_k + b + u_k) - \alpha\overbrace{(Ax_* + b)}^{0}$$

$$e_{k+1} = e_k - \alpha(Ae_k + u_k)$$

$$= (I - \alpha A)e_k - \alpha u_k$$

$$e_1 = (I - \alpha A)e_0 - \alpha u_0$$

$$e_2 = (I - \alpha A)e_1 - \alpha u_1 = (I - \alpha A)^2 e_0 - (I - \alpha A)\alpha u_0 - \alpha u_1$$

$\vdots$

$$e_{k+1} = (I - \alpha A)^k e_0 - \alpha\sum_{j=0}^{k}(I - \alpha A)^{k-j} u_j$$

$u_j \searrow 0$

ex: $u_j = C^{-j}$

$J_k \sim \text{Uniform}(\{1, \dots, n\})$

$x_{k+1} = x_k - \alpha \nabla f_J(x)$

$$u_k = \underbrace{\nabla f_{J_k}(x_k) - \nabla f(x_k)}_{\text{deviation from mean (variance)}}$$

Idea: Sample $J \subset \{1, 2, \dots, n\}$ $|J| = b$ iid. unif. at random

$$p_k = \frac{1}{b} \sum_{j \in J} \nabla f_j(x_k) \qquad \mathbb{E}(p_k) = \frac{1}{b} \sum_{j \in J} \underbrace{\mathbb{E}(\nabla f_J(x_k))}_{\nabla f(x_k)} = \nabla f(x_k)$$

In practice $m = 1, 2, \dots$

for $s = 1, 2, \dots, n/b$ batch size

$p_k = 0$

for $j = (s-1) \cdot b + 1, \dots, s \cdot b$

$p_k \mathrel{+}= \nabla f_j(x_k)$

mini-batch

$x_{k+1} = x_k - \alpha p_k / b$

epoch

$$\frac{2}{\alpha k}\left[\qquad \right]$$