我们要证 $E[F(w_T)] - F^*$ 的收敛上界

首先根据 L-smooth 条件，有

$$E[F(w_T)] - F^* \le \frac{L}{2} E\|w_T - w^*\|^2$$

我们用 t 代替 T，即

$$E[F(w_t)] - F^* \le \frac{L}{2} E\| w_t - w^*\|^2$$

用 Δ_t 代替 $\|w_t - w^*\|^2$，则

$$\Delta_t = E\| w_t - w^*\|^2$$

也就是我们要找到 Δ_t 的上界

这篇论文找 Δ_t 的上界的方式是使用一个归纳证明法的方式，也就是要找出 $\Delta_t \le f(t)$，使得其对于 $\Delta_{t+1} \le f(t+1)$ 时也成立

得到 $\Delta_t \le \frac{v}{\gamma + t}$，$v = \max\left\{\frac{\beta^2 B}{\beta\mu - 1}, (\gamma+1)\Delta_1\right\}$.

其中 $\beta > \frac{1}{\mu}$ 且 $\gamma > 0$.

~~v是后面~~ v 的值是后面决定的

根据引理1、2、3 我们可以得到

$$\Delta_{t+1} \le (1 - \eta_t \mu)\Delta_t + \eta_t^2 B$$，其中 $B = \sum_{k=1}^{N} p_k^2 \sigma_k^2 + 6L\Gamma + 8(E-1)^2 G^2$

则 $\Delta_{t+1} \le (1 - \eta_t\mu)\Delta_t + \eta_t^2 B \longrightarrow$ 令 $\eta_t = \frac{\beta}{t+\gamma}$

$$\le \left(1 - \frac{\beta\mu}{t+\gamma}\right)\cancel{\Delta_t}\frac{v}{t+\gamma} + \frac{\beta^2 B}{(t+\gamma)^2}$$

$$= \frac{t+\gamma-1}{(t+\gamma)^2} v + \left[\frac{\beta^2 B}{(t+\gamma)^2} - \frac{\beta\mu - 1}{(t+\gamma)^2} v\right]$$

$$= \frac{\left[(t+\gamma-1) - (\beta\mu - 1)\right] v + \beta^2 B}{(t+\gamma)^2}$$

令 $v = \max\left\{\frac{\beta^2 B}{\beta\mu - 1}, (r+1)\Delta_1\right\}$ 是为了消项

$$= \frac{(t+r-1)v}{(t+r)^2} \le \frac{(t+r)v}{(t+r)^2} \le \frac{v}{t+r} \le \frac{v}{t+r+1}$$

则我们有，令 $\beta = \frac{2}{\mu}$ ~~可得~~， $k = \frac{L}{\mu}$ 可得

$$E[F(w_t)] - F^* \le \frac{L}{2} E\|w_t - w^*\|^2$$

$$= \cancel{\frac{\Delta}{2}} \le \frac{L}{2}\cdot\frac{v}{r+t}$$

$$\le \frac{L}{2}\cdot\frac{1}{r+t}\cdot\left(\frac{\beta^2 B}{\beta\mu-1} + (r+1)\Delta_1\right)$$

$$= \frac{L}{2}\cdot\frac{1}{r+t}\cdot\left(\frac{4B}{\mu^2} + (r+1)\Delta_1\right)$$

$$= \frac{k}{r+t}\left(\frac{2B}{\mu} + \frac{\mu(r+1)}{2}\Delta_1\right)$$