Summary for single convex funciton optimization \boldsymbol{f}

B stands for B-Lipschitz of f, R is the diameter of X, L stands for L-smoothness, μ stands for μ -strongly convexity, and $\kappa:=L/\mu$.

name	γ_t	condition	last/average iterate convergence	convergence rate w.r.t ${\it T}$	ϵ optimal iteration
GD / PGD	$rac{R}{B\sqrt{T}}$	В	Avg	$\mathcal{O}(1/\sqrt{T})$	$\mathcal{O}(1/arepsilon^2)$
GD / PGD / Mirror GD	$rac{1}{L}$	L	Last	$\mathcal{O}(1/T)$	$\mathcal{O}(1/arepsilon)$
Nesterov Accelerated GD	Addaptive	L	Last	$\mathcal{O}(1/T^2)$	$\mathcal{O}(1/\sqrt{arepsilon})$
GD / PGD	$\frac{1}{L}$	L , μ	Last	$\mathcal{O}(\left(1-rac{\mu}{L} ight)^T)$	$\mathcal{O}(\kappa \ln(rac{1}{arepsilon}))$
GD	$\frac{1}{L}$	L , μ -PL	Last	$\mathcal{O}(\left(1-rac{\mu}{L} ight)^T)$	$\mathcal{O}(\kappa \ln(rac{1}{arepsilon}))$
Coorinate GD	$rac{1}{L_i}$	L , μ	Last	$\mathcal{O}(\left(1-rac{\mu}{d\overline{L}} ight)^T)$	$\mathcal{O}(d\kappa\ln(rac{1}{arepsilon}))$
Sub GD / Mirror GD	$rac{R}{B\sqrt{T}}$	B	Avg	$\mathcal{O}(rac{1}{\sqrt{T}})$	$\mathcal{O}(1/arepsilon^2)$
Sub GD	$rac{f(x_t) - f^*}{\ g(x_t)\ _2^2}, \ rac{2}{\mu(t+1)}$	B , μ	Avg	$\mathcal{O}(rac{1}{T})$	$\mathcal{O}(1/arepsilon)$
Frank-Wolfe	$rac{2}{t+2}$	L	Last	$\mathcal{O}(rac{1}{T})$	$\mathcal{O}(1/arepsilon)$
Newton	Auto	Hessian B - Lipschitz, μ	Last	$\mathcal{O}(\left(rac{1}{2} ight)^{2^T-1})$	$\mathcal{O}(\log\log(1/arepsilon)$
Newton	Auto	L , μ	Last	$\mathcal{O}\left((1-rac{\mu^2}{L^2})^t ight)$	$\mathcal{O}(\kappa^2 \ln(rac{1}{arepsilon}))$

Summary for stochastic funciton optimization of $F=\mathbb{E}[f_{\xi}].$

name	γ_t	condition	last/average iterate convergence	convergence rate w.r.t T	ϵ optimal iteration of f or $ abla f $
SGD	$rac{R}{B\sqrt{T}}$	В	Avg	$\mathcal{O}(rac{1}{\sqrt{T}})$	$\mathcal{O}(rac{1}{\epsilon^2})$
SGD	$rac{1}{2\mu}$	B , μ	Last	$\mathcal{O}(rac{1}{T})$	$\mathcal{O}(rac{1}{\epsilon})$
SGD (!convex)	$\min\left\{rac{1}{L},rac{\gamma}{\sigma\sqrt{T}} ight\}$	$L_{,} \ ext{var}[abla f_{\xi}] \leq \sigma$	Last	$\mathcal{O}(rac{1}{T^{1/4}})$	$\mathcal{O}(rac{1}{\epsilon^4})$

Summary for stochastic funciton optimization of finite sum $F=\sum_i f_i/n$.

name	γ_t	condition	Туре	Error w.r.t T	ϵ -iteration	comp/iter	ε-cost
Full GD	$rac{1}{\overline{L_i}}$	$\overline{L_i}$, μ	Last	$\mathcal{O}(\left(1-rac{\mu}{\overline{L_i}} ight)^T)$	$\mathcal{O}(\kappa \ln(rac{1}{arepsilon}))$	$\mathcal{O}(n)$	$\mathcal{O}(n\kappa\ln(rac{1}{arepsilon}))$
SAG/SAGA	$rac{1}{16L}$	L_i	Last	$\mathcal{O}((1-\min\{rac{C}{n},rac{C'}{\kappa}\})^T)$	$\mathcal{O}((n+\kappa)\ln(rac{1}{\epsilon}))$	$\mathcal{O}(1)$	$\mathcal{O}((n+\kappa)\ln(rac{1}{\epsilon}))$
SVRG	$< \frac{1}{2L}$	L_i	Last	$\mathcal{O}(ho^T))$	$\mathcal{O}((n+\kappa)\ln(rac{1}{\epsilon}))$	$\mathcal{O}(2)$	$\mathcal{O}((n+\kappa)\ln(rac{1}{\epsilon}))$
SPIDER (!convex)	$<\frac{1}{2L}$	Avg- L , $ ext{var}[abla f_{\xi}] {\leq} \sigma^2$	Avg	$\mathcal{O}(rac{1}{\sqrt{T}})$	$\mathcal{O}(rac{1}{\epsilon^2})$	$\mathcal{O}(rac{\sigma}{\epsilon})$	$\mathcal{O}(rac{\sigma}{\epsilon^3})$