Stochastic Gradient Descent from Convex Optimization Perspective

compiled by D.Gueorguiev, 3/29/2025

# Gradient Descent

Consider unconstrained smooth convex optimization

That is, is convex and differentiable with . Denote the optimal criterion value by and a solution by .

Gradient descent algorithm

choose initial point

repeat

stop when certain stopping criterion is satisfied

A grid with colored lines

Description automatically generated

A grid with colored lines and dots

Description automatically generated

//TODO: finish the Gradient Descent section

# The Sub-Gradient Method

//TODO: finish the Sub-Gradient Method

# Proximal Gradient Descent

//TODO: finish the Proximal Gradient Descent

# Stochastic Gradient Descent

//TODO: finish the Stochastic Gradient Descent

# References

[1] [Gradient Descent, Convex Optimization 10-725, CMU, slides, Ryan Tibshirani](https://github.com/dimitarpg13/statistical_learning_and_kernel_methods/blob/main/literature/articles/gradient_descent/grad-descent_Ryan_Tibshirani_slides.pdf)

[2] [Sub-Gradient Method, Convex Optimization 10-725, CMU, slides, Ryan Tibshirani](https://github.com/dimitarpg13/statistical_learning_and_kernel_methods/blob/main/literature/articles/gradient_descent/sub-gradient-method-Ryan_Tibshirani_slides.pdf)

[3] [Proximal Gradient Descent, Convex Optimization 10-725, CMU, slides, Ryan Tibshirani](https://github.com/dimitarpg13/statistical_learning_and_kernel_methods/blob/main/literature/articles/gradient_descent/proximal-gradient-descent_Ryan_Tibshirani_slides.pdf)

[4] [Stochastic Gradient Descent, Convex Optimization 10-725, CMU, slides, Ryan Tibshirani](https://github.com/dimitarpg13/statistical_learning_and_kernel_methods/blob/main/literature/articles/gradient_descent/stochastic-gradient-descent_Ryan_Tibshirani_slides.pdf)