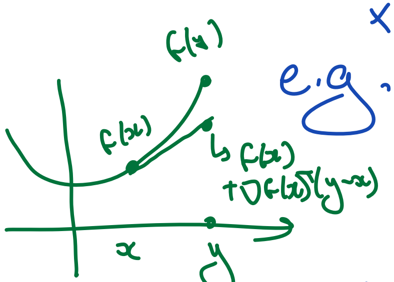
Reminder: Please fill out course evaluations!

Review:

- optimality: • local / global
  - • strict minima

- Taylor's theorem
  - Gradient
  - descent direction

- Convex functions:
  - def'n
  - strict convexity
  - Equivalent cond'ns:



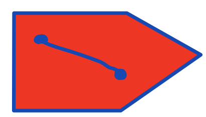
e.g.

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y$$

- $\nabla^2 f(x) \succeq 0 \quad \forall x$

PSD → def'n of PSD

- Convex Sets



  - × def'n
  - × properties

    e.g. C_1, C_2 convex $\Rightarrow C_1 \cap C_2$ convex.

  - × Convexity Theorems (Lectures 2 & 3)
  - * Optimality Theorems (Lecture 5, 6)

- Gradient descent
  - algorithm
  - conv. theorems

(F is Lipschitz convex, differentiable)

(F is smooth, convex, diff.)

Lecture 17

(F is strongly convex, $\|\nabla^2 F\| \leq L$)

- Steepest descent
- Projected GD (for solving constrained opt. problems)
  × projections
- Picking step-size:
  - Backtracking line search
    ↳ theorem
  - Constant acc. to conv. theorems.

Newton's method:

- Algorithm
- Conv. Theorem
- Tradeoffs

- Accelerating GD
  - Polyak Momentum
  - Nesterov Acceleration.

(Algorithms + conv. guarantees in terms of condition number).

- Conjugate Gradient
  - Conjugate directions method for linear systems (conv. theorem)

- Strong Convexity
- PL Condition
- Nonconvex constraints

- Idea about eng applications.