Lecture 17

Review for Midterm

David Gleich
CS 520
Purdue
Spring 2026

If you are interested in ML & Optimization … try and read about …

- Fisher Information Matrix **F**

$$E[\mathbf{g}\mathbf{g}^T]$$

- Hessian **H**

For many ML / ML-like objectives, the Fisher matrix gives Hessian-like and sometimes Hessian-approximate information that's much easier to get.

- AdaGrad
- ADAM

Conditions

- No electronics
- No talking
- No notes
- No scratch paper
- No… anything other than a pen/pencil.
- You can have some type of liquid to consume

Focus on concepts and derivations

~ Half of the test is *conceptual*: "show what this definition means"
e.g. Show the vertices of this polyhedra

~ Half of the test is *derivational*: "show that this case has property X"
e.g. Show that Bp = -g is a descent direction if B is spd.

Topics we have covered

- Sequences & Convergence
- Optimization Software
- Types of minimizers
- Taylor's Theorem
- Multi-D Taylor
- Convexity
- Least squares
- Null-space method
- Linear equality constraints
- Linear inequality constraints
- Lagrange multipliers
- Descent directions
- Optimality & linear inequalities
- Steepest descent / gradient descent with exact line search
- Linear programs
- KKT conditions
- Vertices of polyhedra
- Basic feasible points
- Simplex method
- Phase-1/Crash
- Duality
- Fund. Theorem of LPs
- Line search algorithms
- Wolfe conditions
- Convergence of line search
- Stochastic gradient descent

Concepts we have covered

- **Sequences & Convergence**
- **Optimization Software**
- **Types of minimizers**
- Taylor's Theorem
- Multi-D Taylor
- **Convexity**
- Least squares
- Null-space method
- **Linear equality constraints**
- **Linear inequality constraints**
- **Lagrange multipliers**
- **Descent directions**
- **Optimality & linear inequalities**
- Steepest descent / gradient descent with exact line search
- **Linear programs**
- **KKT conditions**
- **Vertices of polyhedra**
- Simplex method
- Phase-1/Crash
- **Duality**
- **Fund. Theorem of LPs**
- **Line search algorithms**
- **Wolfe conditions**
- Convergence of line search
- Stochastic Gradient Descent

Derivations we have covered

- **Sequences & Convergence**
- Optimization Software
- Types of minimizers
- **Taylor's Theorem**
- **Multi-D Taylor**
- Convexity
- Least squares
- **Null-space method**
- Linear equality constraints
- Linear inequality constraints
- Lagrange multipliers
- Descent directions
- **Optimality & linear inequalities**
- **Steepest descent / gradient descent with exact line search**
- Linear programs
- **KKT conditions**
- Vertices of polyhedra
- Simplex method
- Phase-1/Crash
- **Duality**
- **Fund. Theorem of LPs**
- Line search algorithms
- **Wolfe conditions**
- **Convergence of line search**
- Stochastic gradient descent

Topics we have covered

- **Sequences & Convergence**
- Optimization Software
- **Types of minimizers**
- **Taylor's Theorem**
- Multi-D Taylor
- **Convexity**
- **Least squares**
- Null-space method
- *Linear equality constraints*
- *Linear inequality constraints*
- Lagrange multipliers
- **Descent directions**
- **Optimality & linear inequalities**
- Steepest descent / gradient descent with exact line search
- **Linear programs**
- **KKT conditions**
- **Vertices of polyhedra**
- **Simplex method**
- Phase-1/Crash
- Duality
- Fund. Theorem of LPs
- **Line search algorithms**
- **Wolfe conditions**
- Convergence of line search
- Stochastic gradient descent

My goal with the exam

- These tend to be hard exams and there isn't enough time.
- My goal is to differentiate how well you've understood both the high level concepts of the course and understand the details of how we implement those concepts in algorithms, theory and software.
- I don't like tricky questions or questions that depend on "remembering" a fact. (This could make it EASIER…) but rather, give you a chance to show what you know and have learned.