Week 1

We recalled the basic concepts from analysis needed for the course (gradient, Hessian, first- and second-order approximations). We then introduced convex sets with examples, and established the first- and second-order characterizations of convexity. We also highlighted the power of convexity: convex sets have separating hyperplanes, and locally optimal solutions of convex functions are globally optimal. (Chapter 3 in Vishnoi's book)

Week 2

We introduced convex programs and defined their duals via the Lagrangian function. We showed that strong Lagrangian duality holds under Slater's condition. As examples, we derived the duals of a linear program and a least-squares problem. (Chapter 5 in Vishnoi's book and Lecture 4-5 on Lau's webpage)

Week 3

We discussed the Karush-Kuhn-Tucker (KKT) conditions and showed that, under Slater's condition, a pair of primal and dual solutions are optimal if and only if they satisfy the KKT conditions. We then turned to first-order methods, starting with a high-level overview of gradient descent. In particular, we presented an algorithm for convex functions with a Lipschitz-continuous gradient. (Chapter 6 in Vishnoi's book)

Week 4

We concluded the analysis of the algorithm for functions with Lipschitz-continuous gradients. Then we began discussing the case of bounded gradients and introduced the concept of regularizers. We stated the algorithm for the special case of minimizing a convex function over the probability simplex, which leads to the exponential gradient descent algorithm. Finally, we discussed basic properties of the Bregman divergence, including the law of cosines, the Pythagorean theorem, and Pinsker's inequality (Chapter 7 in Vishnoi's book).

Week 5

We gave a complete analysis of the exponential gradient descent algorithm. We then extended this idea to more general functions and introduced the mirror descent algorithm. We observed that the proof carries over with only minor modifications, and that the same reasoning applies to the Multiplicative Weights Update method. (Chapter 7 in Vishnoi's book)

Week 6

We began discussing Newton's method, first focusing on the one-dimensional setting. We showed that the algorithm converges quadratically, provided the starting point is sufficiently close to a root of the function. We then defined the Newton step in the multivariate setting and applied it to the minimization of a convex function. Finally, we introduced the Newton-Euclidean condition and examined when it can be effectively applied. (Chapter 9 is Vishnoi's book)

Week 7

We defined the local norm, formulated the Newton local condition, and analyzed Newton's method under this condition. We then sketched the general idea of primal path-following interior-point methods.

Week 8

Assuming that we start with an initial point x_0 and a value η_0 such that $\left\|n_{\eta_0}(x_0)\right\|_{x_0} \leq \frac{1}{6}$, we proved that this property is maintained throughout the algorithm. We then showed that, after performing two further Newton steps from x_T, the objective value at the resulting $\hat{x}$ is close to the optimum value. (Chapter 10 in Vishnoi's book)

Week 8

We finished the description of the interior point method: we discussed how to find a central path through a point x' inside the polytope, how to follow this path in reverse until we get close to the analytic center, and then how to change to the original central path. Then we showed how to reduce optimization to deciding feasibility, and then deciding feasibility to separation.

Week 10

We discussed the main steps of the ellipsoid method and showed that for $0-1$ polytopes with an integer cost function, rounding each coordinate of the resulting vector yields an optimal solution. (Chapter 12 of Vishnoi's book)