# Numerical Optimization Final Exam

Fan Zhang and Xiangyu Yang

December 30, 2020

**1** (15 = 5+5+5 points) Consider a linear system of equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{x} \in \mathbb{R}^n$ and that $\boldsymbol{A}$ is positive definite. This is equivalent to minimizing a quadratic function $\phi(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}$.

(i) Show that if a set of nonzero vectors $\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_m\} \in \mathbb{R}^n$ $(m < n)$ are $\boldsymbol{A}$-conjugate, then they are linearly independent.

(ii) Suppose we have an initial point $\boldsymbol{x}_0$ and initial search direction $\boldsymbol{p}_0 = -\nabla\phi(\boldsymbol{x}_0)$. What is the exact line-search stepsize along $\boldsymbol{p}_0$?

(iii) Suppose we define the new direction as $\boldsymbol{p}_1 = -\boldsymbol{r}_1 + \beta\boldsymbol{p}_0$ (where $\boldsymbol{r}_1$ is the residual at $\boldsymbol{x} = \boldsymbol{x}_1$) and require $\boldsymbol{p}_0, \boldsymbol{p}_1$ are $\boldsymbol{A}$-conjugate. What is the value for $\beta$?

**2** (10 points) Find the projection of $\boldsymbol{y} \in \mathbb{R}^n$ onto the half-hyperspace $\{\boldsymbol{x} \mid \boldsymbol{a}^T\boldsymbol{x} \leq \boldsymbol{b}\}$ with $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{a} \neq \mathbf{0}$. In other words, solve the following Euclidean projection problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 \quad \text{s.t. } \boldsymbol{a}^T\boldsymbol{x} \leq b. \tag{0.4}$$

**3** (15 points) Consider the inequality constrained strictly convex quadratic programming (QP) problem

$$
\begin{aligned}
\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \tfrac{1}{2}\boldsymbol{x}^T\boldsymbol{H}\boldsymbol{x} + \boldsymbol{g}^T\boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{a}_i^T\boldsymbol{x} + b_i = 0, \quad i = 1,\dots,m \\
& \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, \quad i = m+1,\dots,t.
\end{aligned}
\tag{0.8}
$$

Suppose $\boldsymbol{x}^*$ is the first-order optimal solution and the active-set at $\boldsymbol{x}^*$ is $\mathcal{A}(\boldsymbol{x}^*) := \{i \in \{m+1,\dots,t\} \mid \boldsymbol{a}_i^T\boldsymbol{x}^* + b_i = 0\}$. Show that $\boldsymbol{d} = \boldsymbol{0}$ is optimal for the following problem

$$
\begin{aligned}
\min_{\boldsymbol{d}\in\mathbb{R}^n} \quad & \tfrac{1}{2}(\boldsymbol{x}^*+\boldsymbol{d})^T\boldsymbol{H}(\boldsymbol{x}^*+\boldsymbol{d}) + \boldsymbol{g}^T(\boldsymbol{x}^*+\boldsymbol{d}) \\
\text{s.t.} \quad & \boldsymbol{a}_i^T(\boldsymbol{x}^*+\boldsymbol{d}) + b_i = 0, \quad i = 1,\dots,m \\
& \boldsymbol{a}_i^T(\boldsymbol{x}^*+\boldsymbol{d}) + b_i \le 0, \quad i \in \mathcal{A}(\boldsymbol{x}^*).
\end{aligned}
\tag{0.9}
$$

**4** (15 = 5 + 10 points) In a quasi-Newton method for solving the unconstrained optimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \; f(\boldsymbol{x}).$$

We use local model

$$m_k(\boldsymbol{d}) = f(\boldsymbol{x}_k) + \nabla f(\boldsymbol{x}_k)^T \boldsymbol{d} + \tfrac{1}{2}\boldsymbol{d}^T \boldsymbol{H}_k \boldsymbol{d}$$

to approximate $f(\boldsymbol{x})$ at $\boldsymbol{x}_k$. The secant equation is obtained by requiring the gradient of $m_k(\boldsymbol{d})$ at $\boldsymbol{x}_{k-1}$ is equivalent to $\nabla f(\boldsymbol{x}_{k-1})$.

(i) Derive the secant equation that must satisfy.

(ii) Suppose you were using a multiple of identity matrix $\alpha \boldsymbol{I}$ to approximate the Hessian matrix (i.e., $\boldsymbol{H}_k = \alpha \boldsymbol{I}$), which may not satisfy the secant equation. Find the $\alpha$ as the least-squares solution of the secant equation. (The least squares solution of a linear equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is the minimizer of $\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$.)

**5** (30 = 10 × 3 points) Consider the unconstrained optimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \; f(\boldsymbol{x}),$$

where $f \in C^2$. In addition, we have the following assumptions on $f$:

(1) $f$ is $L$-smooth.

(2) $f$ is bounded below over $\boldsymbol{x} \in \mathbb{R}^n$.

(i) Suppose $\boldsymbol{d}_k \in \mathbb{R}^n$ is a descent direction at $\boldsymbol{x}_k$. Show that

$$f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) \leq f(\boldsymbol{x}_k)$$

for sufficiently small stepsize $\alpha_k > 0$.

(ii) The Armijo line search condition is

$$f(\boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k) \leq f(\boldsymbol{x}_k) + \eta \alpha_k \nabla f(\boldsymbol{x}_k)^T \boldsymbol{d}_k$$

with $\eta \in (0, 1)$. Show that for a sufficiently small stepsize, this condition must hold (provide the expression of this stepsize).

(iii) Now let $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$ with $\boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)$, show that

$$\|\nabla f(\boldsymbol{x}_k)\|_2 \to 0.$$

**6** (15 points) Consider the unconstrained optimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad f(\boldsymbol{x}).$$

The trust region subproblem subproblem is given by

$$\min_{\boldsymbol{d}\in\mathbb{R}^n} \nabla f(\boldsymbol{x}_k)^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}\boldsymbol{d} \qquad \text{s.t.} \|\boldsymbol{d}\|_2 \leq \Delta_k. \tag{0.17}$$

Derive the Cauchy-point of this subproblem.