Study Homework Questions 1 Numerical Optimization Fall 2023

Problem 1.1

Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{S}$ be a subspace with dimension k and basis $\{\phi_j\}_{j=1}^k$ so that

$$\forall\, s \in \mathcal{S}, \quad s = \phi_1\gamma_1 + \ldots + \phi_k\gamma_k$$

for a unique $c \in \mathbb{R}^k$ with $e_i^T c = \gamma_i$.

Recall that the solution to the subspace approximation problem given by

$$\forall\, h \in \mathcal{H}, \quad h = h_{\mathcal{S}} + h_{\mathcal{S}^\perp}$$

$$h_{\mathcal{S}} = \phi_1\gamma_1^* + \cdots + \phi_k\gamma_k^* = \underset{s\in\mathcal{S}}{\text{argmin}}\|h - s\|^2$$

$$c^* = \begin{pmatrix} \gamma_1^* \\ \vdots \\ \gamma_k^* \end{pmatrix} \in \mathbb{R}^k, \quad g = \begin{pmatrix} \langle h, \phi_1 \rangle \\ \vdots \\ \langle h, \phi_k \rangle \end{pmatrix} \in \mathbb{R}^k$$

$$Gc^* = g$$

where $G \in \mathbb{R}^{k\times k}$ is the Gram matrix defined by the basis vector $\{\phi_1, \ldots, \phi_k\}$.

(1.1.a) Show that the Gram matrix G is nonsingular.

(1.1.b) Under what circumstances can $c^* = 0 \in \mathbb{R}^k$?

(1.1.c) Suppose you want to find $h_{\mathcal{S}^\perp}$ where $\mathcal{S}$ has dimension k but $\mathcal{H}$ is infinite dimensional. Write the optimization problem that determines $h_{\mathcal{S}^\perp}$ using explicit functional constraints and explain how you would compute the answer.

Problem 1.2

(Luenberger, Optimization by Vector Space Methods, 1969, pp. 66-67.)

The angular velocity, $\omega(t)$, and angular position, $\theta(t)$, of a DC motor shaft driven by a current source $u(t)$ are governed by the differential equations

$$\dot{\omega}(t) = -\omega(t) + u(t), \quad \omega(0) = 0 \tag{1}$$

$$\dot{\theta}(t) = \omega(t), \quad \theta(0) = 0 \tag{2}$$

where the initial condition $[\omega(0), \theta(0)] = [0, 0]$ is an at rest position at 0 angle.

Assume $u,\ \omega,\ \theta \in \mathcal{L}^2[0, 1]$ with

$$\langle z, q\rangle = \int_0^1 z(\tau)q(\tau)d\tau, \quad \|z\|^2 = \int_0^1 z^2(\tau)d\tau.$$

The optimal control problem is to determine the current profile control action with minimum energy $\|u\|^2$ so that the system moves from its initial resting state $[\omega(0), \theta(0)] = [0, 0]$ to final resting state $[\omega(1), \theta(1)] = [0, 1]$ in one second.

Recall, that the solution of the initial value problem

$$\dot{f}(t) = -\alpha f(t) + g(t), \quad f(0) = \phi$$

is

$$f(t) = \phi e^{-\alpha t} + \int_0^t e^{\alpha(\tau - t)}\ g(\tau)\ d\tau.$$

(1.2.a) Find expressions for $\omega(1)$ and $\theta(1)$ in terms of $u(t)$.

(1.2.b) Show that the optimal control problem can be expressed as

$$\min_{u\in\mathcal{L}^2[0,1]} \quad \|u\|^2$$

$$\text{subject to } \ \langle y_1, u\rangle {=} \gamma_1 \ \text{ and } \ \langle y_2, u\rangle {=} \gamma_2$$

for some functions $y_1(t),\ y_2(t) \in \mathcal{L}^2[0, 1]$ and constants $\gamma_1,\ \ \gamma_2 \in \mathbb{R}$.

(1.2.c) Find the optimal control $u(t)$ using Hilbert space techniques discussed in the notes.

Problem 1.3

Recall the "chord" definition of a convex function $f : \mathbb{R} \to \mathbb{R}$

Definition 1.3.1. $f : \mathbb{R} \to \mathbb{R} : x \mapsto f(x)$ is convex on $[a, b]$ if $\forall x_0,\ x_1 \in [a, b]$

$$f(\theta x_1 + (1 - \theta)x_0) \leq \ell(\theta;\ x_0,\ x_1)$$

$$\ell(\theta;\ x_0,\ x_1) = \theta f(x_1) + (1 - \theta)f(x_0), \quad 0 \leq \theta \leq 1$$

$$\begin{aligned}\ell(x;\ x_0,\ x_1) &= f(x_0) + \theta(f(x_1) - f(x_0)) = f(x_0) + \theta\Delta x f[x_0, x_1] \\ &= f(x_0) + (x - x_0)f[x_0, x_1]\end{aligned}$$

$$0 \leq \theta \leq 1,\ \Delta x = x_1 - x_0,\ x = x_0 + \phi\Delta x,\ f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

Prove the following theorem:

Theorem 1.3.1.

1. *If $f \in \mathcal{C}^1$ then the tangent condition*

$$f(x) \geq t(x,\ x_0) = f(x_0) + (x - x_0)f'(x_0), \quad \forall\ x,\ x_0 \in [a,b]$$

   is equivalent to the chord condition given in Definition 1.3.1.

2. *If $f \in \mathcal{C}^1$ and convex then*

$$0 \leq (x - y)(f'(x) - f'(y)), \quad \forall\ x,\ y \in [a,b]$$

3. *If $f \in \mathcal{C}^2$ then the condition that $f''(x) \geq 0$ $\forall x \in [a,b]$ is equivalent to the tangent condition above and the chord condition given in Definition 1.3.1.*

Problem 1.4

Consider the minimization problem

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

where $f(x) = \frac{1}{2}x^T Ax - x^T b$, $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite, and $b \in \mathbb{R}^n$.

(1.4.a) Show that $\forall 0 \leq \beta \leq 1$

$$\beta f(x) \geq f(\beta x)$$

(1.4.b) Show that $f(x)$ is a convex function.

Problem 1.5

Suppose $A \in \mathbb{R}^{n\times n}$ is a **symmetric positive semidefinite** matrix and $f(x) = 0.5x^T Ax - x^T b$ with $b \in \mathbb{R}^n$ and $b \in \mathcal{R}(A)$. Show that Steepest Descent will converge to an unconstrained minimizer of $f(x)$ for any x_0 such that $Ax_0 \neq 0$.

Hint: Find a smaller, symmetric positive definite linear system and use the fact that steepest descent converges on a symmetric positive definite system.

Problem 1.6

(Problem 17 on page 259 of Luenberger and Ye 3rd Ed.)

Suppose the method of Steepest Descent is used to minimize

$$f(x) = \frac{1}{2}(x - x^*)^T A(x - x^*)$$

and the stepsize α_k is not determined to minimize $f(x_k + \alpha r_k)$ as a function of the scalar α but instead only satisfies

$$\frac{E(x_k) - E(x_{k+1})}{E(x_k)} \geq \beta \frac{E(x_k) - \bar{E}}{E(x_k)}$$

where

$$f(x_k) = E(x_k) = 0.5\|x_k - x_*\|_A^2$$

for some $0 < \beta < 1$, where $\bar{E}$ is the value that corresponds to the best α_k, i.e., the usual minimizer. Find the best estimate for the convergence rate of the algorithm.

Problem 1.7

(Problem 21 on page 260 of Luenberger and Ye 3rd Ed.)

Let $x \in \mathbb{R}^2$ with elements ξ_1 and ξ_2. Consider the cost function

$$f(x) = \xi_1^2 + \xi_2^2 + \xi_1\xi_2 - 3\xi_1.$$

(1.7.a) Find an unconstrained local minimizer.

(1.7.b) What is the rate of convergence for Steepest Descent applied to the unconstrained problem?

(1.7.c) Is the local minimizer also a global minimizer?

(1.7.d) Suppose the constraints $\xi_1 \geq 0$ and $\xi_2 \geq 0$ are added. Can a minimizer still be determined? If so, what is it?

Problem 1.8

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix, $C \in \mathbb{R}^{n \times n}$ be a symmetric nonsingular matrix, and $b \in \mathbb{R}^n$ be a vector. The matrix $M = C^2$ is therefore symmetric positive definite. Also, let $\tilde{A} = C^{-1}AC^{-1}$ and $\tilde{b} = C^{-1}b$.

The preconditioned Steepest Descent algorithm to solve $Ax = b$ is:

A, M are symmetric positive definite
x_0 arbitrary; $r_0 = b - Ax_0$; solve $Mz_0 = r_0$

do $k = 0, 1, \ldots$ until convergence

$$w_k = Az_k$$
$$\alpha_k = \frac{z_k^T r_k}{z_k^T w_k}$$
$$x_{k+1} \leftarrow x_k + z_k\alpha_k$$

$r_{k+1} \leftarrow r_k - w_k\alpha_k$
solve $Mz_{k+1} = r_{k+1}$

end

The Steepest Descent algorithm to solve $\tilde{A}\tilde{x} = \tilde{b}$ is:

$\tilde{A}$ is symmetric positive definite
$\tilde{x}_0$ arbitrary; $\tilde{r}_0 = \tilde{b} - \tilde{A}\tilde{x}_0$; $\tilde{v}_0 = \tilde{A}\tilde{r}_0$

do $k = 0, 1, \ \ldots$ until convergence

$\tilde{\alpha}_k = \frac{\tilde{r}_k^T\tilde{r}_k}{\tilde{r}_k^T\tilde{v}_k}$
$\tilde{x}_{k+1} \leftarrow \tilde{x}_k + \tilde{r}_k\tilde{\alpha}_k$
$\tilde{r}_{k+1} \leftarrow \tilde{r}_k - \tilde{v}_k\tilde{\alpha}_k$
$\tilde{v}_{k+1} \leftarrow \tilde{A}\tilde{r}_{k+1}$

end

Show that given the appropriate consistency between initial guesses the preconditioned steepest descent recurrences to solve $Ax = b$ can be derived from the steepest descent recurrences to solve $\tilde{A}\tilde{x} = \tilde{b}$.

Problem 1.9

1.9.a

The following lemma is a classic result for convex and concave functions on $\mathbb{R}$.

Lemma (Jensen's Inequality)**.** *Suppose a function $f : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$, scalars $0 \leq \lambda_k \leq 1$ such that $\sum_{k=1}^n \lambda_k = 1$, and scalars $\xi_k \in \mathcal{D}$, $k = 1, \ldots, n$ are given.*

If $f(\xi)$ is convex on $\mathcal{D}$ then

$$f\left(\sum_{k=1}^n \lambda_k\xi_k\right) \leq \sum_{k=1}^n \lambda_k f(\xi_k)$$

and $f(\xi)$ is concave on $\mathcal{D}$ then

$$f\left(\sum_{k=1}^n \lambda_k\xi_k\right) \geq \sum_{k=1}^n \lambda_k f(\xi_k).$$

Show that if $0 \geq \alpha_k \in \mathcal{R}$ and $0 < p_k \in \mathcal{Z}$ with $\sum_{k=1}^n 1/p_k = 1$, $k = 1, \ldots, n$ are given then

$$\prod_{k=1}^n \alpha_k \leq \sum_{k=1}^n \frac{1}{p_k}\ \alpha_k^{p_k}.$$

This is Generalized Young's Inequality. Young's Inequality is with $n = 2$.

1.9.b

Use the Generalized Young's Inequality to prove the inequality relating the arithmetic mean to the geometric mean of positive real numbers $0 < \xi_k < \infty$, $k = 1, \ldots, n$,

$$\left(\prod_{k=1}^{n} \xi_k \right)^{1/n} \leq \frac{1}{n} \sum_{k=1}^{n} \xi_k.$$

Problem 1.10

Prove the following

Lemma (Hoelder's Inequality)**.** *If* $a,\ b \in \mathbb{R}^n$ *and positive integers* p *and* q *satisfy* $p^{-1}+q^{-1} = 1$, *equivalently* $p = q/(q-1)$ *or* $q = p/(p-1)$, *then*

$$a^T b \leq |a^T b| \leq \|a\|_p \|b\|_q.$$

Problem 1.11

Prove the following

Lemma (Minkowski's Inequality)**.** *If* $x,\ y \in \mathbb{R}^n$ *and* $p > 0$ *is an integer*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$