Homework 4 Foundations of Computational Math 2 Spring 2021

These study questions relate to the material optimization-based iterative methods for linear systems defined by symmetric positive definite matrices.

Problem 4.1

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 $b \in \mathbb{R}^n$.

(4.1.a) Show that $\forall 0 \le \beta \le 1$

$$\beta f(x) \ge f(\beta x)$$

(4.1.b) Suppose $x_0 \in \mathbb{R}^n$, $x_1 \in \mathbb{R}^n$ and $x_0 \ne x_1$. Show that $f(x)$ is a convex function.

Problem 4.2

Let $A = Q\Lambda Q^T$ be a symmetric positive definite matrix where Q is an orthogonal matrix and Λ is a diagonal matrix whose diagonal elements are positive and also are the eigenvalues of A. Define

$$\tilde{x} = Q^T x \quad \text{and} \quad \tilde{b} = Q^T b$$
$$Ax = b \quad \text{and} \quad \Lambda\tilde{x} = \tilde{b}$$

Given x_0 and $\tilde{x}_0$, define the sequence x_k as the sequence of vectors produced by steepest descent applied to $Ax = b$ and the sequence $\tilde{x}_k$ as the sequence of vectors produced by steepest descent applied to $\Lambda\tilde{x} = \tilde{b}$.

Let $e_k = x_k - x$ and $\tilde{e}_k = \tilde{x}_k - \tilde{x}$. Show that if $\tilde{x}_0 = Q^T x_0$ then

$$\|e_k\|_2 = \|\tilde{e}_k\|_2, \quad k > 0$$

Problem 4.3

Let $A \in \mathbb{R}^{n\times n}$ be symmetric postive definite with an eigendecompositon $A = Q\Lambda Q^T$ with $Q \in \mathbb{R}^{n\times n}$ and orthogonal matrix, i.e., $Q^TQ = QQ^T = I$, and $\Lambda \in \mathbb{R}^{n\times n}$ a diagonal matrix with positive diagonal elements $\lambda_i = e_i^T\Lambda e_i > 0$.

Consider the two systems $Ax = b$ and $\Lambda\tilde{x} = \tilde{b}$ with $Q\tilde{x} = x$ and $Q\tilde{b} = b$. The iterations defined by applying Steepest Descent (SD) to each are

$$x_{k+1} = x_k + \alpha_k r_k, \quad r_k = b - Ax_k, \quad \alpha_k = \frac{r_k^T r_k}{r_k^T A r_k}$$

$$\tilde{x}_{k+1} = \tilde{x}_k + \tilde{\alpha}_k \tilde{r}_k, \quad \tilde{r}_k = \tilde{b} - \Lambda\tilde{x}_k, \quad \tilde{\alpha}_k = \frac{\tilde{r}_k^T \tilde{r}_k}{\tilde{r}_k^T \Lambda \tilde{r}_k}$$

given x_0 and $Q\tilde{x}_0 = x_0$. The elements of the vectors with the tildes are the coefficients of the corresponding vectors without the tildes with respect to the basis of eigenvectors given by the columns of Q.

(4.3.a) Show that $\alpha_k = \tilde{\alpha}_k$ and that

$$\alpha_k^{-1} = \tilde{\alpha}_k^{-1} = \sum_{i=1}^{n} \gamma_i \lambda_i, \quad \gamma_i \geq 0, \quad \sum_{i=1}^{n} \gamma_i = 1.$$

(4.3.b) Any $x_0 \in \mathbb{R}^n$ can be corrected to $A^{-1}b$ by

$$A^{-1}b = x_0 + c_0, \quad c_0 = A^{-1}(b - Ax_0) = A^{-1}r_0.$$

Consider applying SD to $Ax = b$. Derive a sufficient condition on A so that for any x_0 convergence to $A^{-1}b$ occurs in one step, i.e.,

$$A^{-1}b = x_1 = x_0 + \alpha_0 r_0.$$

(4.3.c) Is the condition also a necessary condition for convergence of SD in one step for any x_0?

Problem 4.4

4.4.a

Consider the iteration:

$$\begin{aligned} y_0 &= 0 \\ y_{i+1} &= y_i + \tilde{\alpha}_i e_{i+1} \\ \tilde{r}_i &= b - Dy_i \\ \tilde{\alpha}_i &= \frac{e_{i+1}^T \tilde{r}_i}{e_{i+1}^T D e_{i+1}} \end{aligned}$$

where $D \in \mathbb{R}^{n \times n}$ is a nonsingular diagonal matrix.

Show that $y_n = y = D^{-1}b$.

4.4.b

Suppose $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix and $Ax = b$. Let $p_1, p_2, \cdots, p_n$ be real vectors that are A-orthogonal, i.e., $< p_i, p_j >= 0$ if $i \neq j$ and $< p_i, p_i >> 0$ where $< w, v >= w^T A v$ is the inner product on $\mathbb{R}^n$ defined by A.

Use the result from the first part of the problem to show that the conjugate direction iteration:

$$\begin{aligned} x_0 &= 0 \\ x_{i+1} &= x_i + \alpha_i p_{i+1} \\ r_i &= b - Ax_i \\ \alpha_i &= \frac{p_{i+1}^T r_i}{p_{i+1}^T A p_{i+1}} \end{aligned}$$

is such that $x_n = x = A^{-1}b$.

Problem 4.5

The conjugate direction iteration (CD) can also be derived from a basis expansion point of view. Let $e_{true} = x - x_0 = A^{-1}b - x_0$ where A is a symmetric positive definite matrix. Let $< w, v >= w^T A v$ be the inner product on $\mathbb{R}^n$ defined by A and $p_1, p_2, \cdots, p_n$ be real vectors that are A-orthonormal, i.e., $< p_i, p_j >= 0$ if $i \neq j$ and $< p_i, p_i >= 1$.

4.5.a Show that any vector can be easily written in terms of a basis that is orthonormal with respect to some inner product and apply this to e_{true} to get

$$e_{true} = p_1 < p_1, e_{true} > + \cdots p_n < p_n, e_{true} > \tag{1}$$

4.5.b Show that for any x_0

$$\begin{aligned} \alpha_i &= p_{i+1}^T r_i \\ x_{i+1} &= x_i + \alpha_i p_{i+1} \\ r_i &= b - Ax_i \end{aligned}$$

is such that $x_n = x = A^{-1}b$.

Hint: Define an iteration based on (1) that yields $x_n = x$ and then show it can be computed via the CD iteration given in this problem.

Problem 4.6

Recall the basic CD/CG properties that hold given the assumption that CG has not converged at step k,

- $x_k = \alpha_0 d_0 + \cdots + \alpha_{k-1} d_{k-1}$ is optimal (inherited from CD), i.e.,

$$\forall x \in x_0 + span[d_0, d_1, \ldots, d_{k-1}], \ \ \|x_k - A^{-1}b\|_A \le \|x - A^{-1}b\|_A$$

- $< d_k,\ d_j >_A = 0 \ i \ne j$ for $0 \le i, j \le k-1$ (inherited from CD).
- $< r_k,\ d_j >= 0$ for $0 \le j \le k-1$ (inherited from CD).
- $< r_k,\ r_j >= 0$ for $0 \le j \le k-1$ (CG-specific).
- $span[d_0, d_1, \ldots, d_k] = span[r_0, r_1, \ldots, r_k]$ (CG-specific).
- $span[r_0, r_1, \ldots, r_k] = span[r_0, Ar_0, \ldots, A^k r_0]$ (CG-specific).

Given the inherited properties prove the three CG-specific properties.

Problem 4.7

4.7.a

Let the cost function $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x) = x^T d + x^T x, \quad \text{where} \quad d = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}$$

where $\delta_1 > 0$, $\delta_2 > 0$ and $\mu = \|d\|_2 > 1$. Consider the problem

$$\min_{x \in \mathbb{R}^2} f(x).$$

(i) Find a minimizer x^*. Is it unique?

(ii) Write the iteration that defines applying the steepest descent algorithm to solve the minimization problem.

(iii) How would you set the stepsize α_k and why?

(iv) Will your choice of α_k yield an algorithm that converges in a finite number of steps?

4.7.b

Now suppose the minimization problem is constrained so that we are only interested in $x \in \mathbb{R}^2$ on the circle of radius 1, i.e., the unit circle

$$\mathcal{S}_1 = \{x \in \mathbb{R}^2 \mid x^T x = 1\}$$

Specifically, we want to solve

$$\min_{x \in \mathcal{S}_1} f(x)$$

(i) Show that this problem can be viewed as an unconstrained minimization problem on $\mathbb{R}$ by writing the cost function over $\mathcal{S}_1$ as a function of a real variable θ.

(ii) Write the iteration that defines applying the steepest descent algorithm to solve the minimization problem over $\mathbb{R}$.

(iii) How would you set the stepsize α_k and why?

(iv) Will your choice of α_k yield an algorithm that converges in a finite number of steps when started at an initial guess $\theta_0 = 0$?