Numerical Optimization with applications: Homework 04

104021601 林俊傑
104021602 吳彥儒
104021615 黃翊軒

November 9, 2016

Exercise 1. *Implement Algorithm 5.2 and use to it solve linear systems in which A is the Hilbert matrix, whose elements are $A_{i,j} = 1/(i + j - 1)$. Set the right-hand-side to $b = (1, 1, ..., 1)^T$ and the initial point to $x_0 = 0$. Try dimensions $n = 5, 8, 12, 20$ and report the number of iterations required to reduce the residual below 10^{-6}.*

Solution. The numbers of iterations as the table below.

n	5	8	12	20
number of iteration	6	19	38	73
condition number	4.766E+05	1.526E+10	1.633E+16	2.596E+18

Observe that the condition number in the case $n = 20$ is greater than the others. By(5.36), the rate of convergence should be less than the others. ◀

Exercise 2. *Show that if the nonzero vectors $p_0, p_1, ..., p_l$ satisfy (5.5), where A is symmetric and positive definite, then these vectors are linearly independent. (This result implies that A has at most n conjugate direction.)*

Proof. Suppose $a_0p_0 + a_1p_1 + ... + a_lp_l = 0$. For any p_j, we have the following argument.

$$\begin{aligned}
0 &= p_j^T A(a_0p_0 + a_1p_1 + ... + a_lp_l) \\
&= a_0(p_j^T Ap_0) + a_1(p_j^T Ap_1) + ... + a_j(p_j^T Ap_j) + ... + a_l(p_j^T Ap_l) \\
&= a_0 \cdot 0 + a_1 \cdot 0 + ... + a_j \cdot (p_j^T Ap_j) + ... + a_l \cdot 0 \\
&= a_j \cdot (p_j^T Ap_j)
\end{aligned}$$

Since A is positive definite, $p_j^T Ap_j > 0$, this implies $a_j = 0$. $\quad \forall j$
Consequently, $p_0, p_1, ..., p_l$ are linearly independent.

□

Exercise 4. *Show that if $f(x)$ is a strictly convex quadratic, then the function $h(\sigma) \stackrel{def}{=} f(x_0 + \sigma_0p_0 + \cdots + \sigma_{k-1}p_{k-1})$ also is a strictly convex quadratic in the variable $\sigma = (\sigma_0, \sigma_1, \cdots, \sigma_{k-1})^T$.*

Proof. By the definition of strictly convex quadratic function, we can assume

$$f(x) = \frac{1}{2}x^T Ax - b^T x,$$

where A is a positive definite symmetric matrix and b is a constant vector. We want prove that $h(\sigma)$ is also a strictly convex quadratic function by showing

$$h(\sigma) = \frac{1}{2}\sigma^T B\sigma - c^T\sigma + d,$$

where B is a positive definite symmetric matrix and c, d are constant vectors. Since $p_i^T Ap_j = 0$ for all

$i \neq j$, we obtain that

$$
\begin{aligned}
h(\sigma) &= f(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) \\
&= \frac{1}{2}(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1})^T A(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) - b^T(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) \\
&= \frac{1}{2} x_0^T A(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) + \frac{1}{2}(\sigma_0 p_0)^T A(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) + \cdots \\
&\quad + \frac{1}{2}(\sigma_{k-1} p_{k-1})^T A(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) - b^T(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) \\
&= \frac{1}{2}(\sigma_0 p_0)^T A(\sigma_0 p_0) + \frac{1}{2}(\sigma_1 p_1)^T A(\sigma_1 p_1) + \cdots + \frac{1}{2}(\sigma_{k-1} p_{k-1})^T A(\sigma_{k-1} p_{k-1}) \\
&\quad + \frac{1}{2} x_0^T A(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) - b^T(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1}) \\
&= \frac{1}{2}\sigma^T B\sigma + \frac{1}{2} x_0^T A P\sigma - b^T P\sigma + \frac{1}{2} x_0^T A x_0 - b^T x_0 \\
&= \frac{1}{2}\sigma^T B\sigma + (\frac{1}{2} x_0^T A P - b^T P)\sigma + \frac{1}{2} x_0^T A x_0 - b^T x_0
\end{aligned}
$$

where $B = \begin{bmatrix} p_0^T A p_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_{k-1}^T A p_{k-1} \end{bmatrix}$ is positive definite sysmmetric matrix and $P = (p_0, p_1, \cdots, p_{k-1})$ and $\sigma = (\sigma_0, \sigma_1, \cdots, \sigma_{k-1})^T$. Hence, $h(\sigma)$ is also a strictly convex quadratic function. ☐

Exercise 7. *Let $\{\lambda_i, v_i\}$ $i = 1, 2, \cdots, n$ be the eigenpairs of the symmetric matrix A. Show that the eigenvalues and eigenvectors of $[I + P_k(A)A]^T A[I + P_k(A)A]$ are $\lambda_i[1 + \lambda_i P_k(\lambda_i)]^2$ and v_i, respectively.*

Proof. We first show that

$$P_k(A)v_i = P_k(\lambda_i)v_i$$

for any polynomials $P_k(x)$ of degree k.

Let $P_k(x) = \sum_{j=0}^{k} a_j x^j$. Then

$$P_k(A)v_i = \sum_{j=0}^{k} a_j A^j v_i = \sum_{j=0}^{k} a_j A^{j-1}(\lambda_i v_i) = \sum_{j=0}^{k} a_j A^{j-2}(\lambda_i^2 v_i) = \cdots = \sum_{j=0}^{k} a_j \lambda_i^j v_i = P_k(\lambda_i)v_i$$

Since $[I + P_k(x)x]$ is a ploynomial, we have

$$[I + P_k(A)A]v_i = [1 + \lambda_i P_k(\lambda_i)]v_i$$

A is symmetric, therefore, $[I + P_k(A)A]^T = [I + P_k(A)A]$

Now,we are ready to compute

$$
\begin{aligned}
[I + P_k(A)A]^T A[I + P_k(A)A]v_i &= [I + P_k(A)A]A[I + P_k(A)A]v_i \\
&= [I + P_k(A)A]A[1 + \lambda_i P_k(\lambda_i)]v_i \\
&= [I + P_k(A)A](Av_i)[1 + \lambda_i P_k(\lambda_i)] \\
&= [I + P_k(A)A]\lambda_i v_i[1 + \lambda_i P_k(\lambda_i)] \\
&= [I + P_k(A)A]v_i \lambda_i[1 + \lambda_i P_k(\lambda_i)] \\
&= [1 + \lambda_i P_k(\lambda_i)]v_i \lambda_i[1 + \lambda_i P_k(\lambda_i)] \\
&= \lambda_i[1 + \lambda_i P_k(\lambda_i)]^2 v_i
\end{aligned}
$$

We conclude that $\{\lambda_i[1 + \lambda_i P_k(\lambda_i)]^2, v_i\}$ $i = 1, 2, \cdots, n$ are the eigenpairs of $[I + P_k(A)A]^T A[I + P_k(A)A]$. ☐