COL726

Major exam

K.Laxman
2018CS50408

6) a) To prove $x^{(0)}, x^{(1)}, x^{(2)} \ldots$ converges to x^* iff $f'(x)$ converges to zero we can use the first order optimality condition for convex functions.

i.e $x^{(k)}$

It states that a point x^* is a minimizer of a convex fn iff $f(x^*) = 0$

Proving it both sides

1) if $x^0, x^1, x^2, \ldots$ converges to x^k then $f'(x^k)$ converges to zero

Given that x^* is unique minimizer of f and f' is twice diff. convex function

$f'(x^*) = 0$

Since $x^{(0)}, x^{(1)}, x^{(2)} \ldots$ converges to x^* (we can use continuity of f & f') to establish $f'(x^k)$ converges to zero

i.e. f is convex & $f'(x^k) \to 0$, we can use continuity of f' s.t. $f(x^*) = 0$ here x^* is unique minimizer of f

Now by contradiction assume that $x^{(0)}, x^{(1)}, x^{(2)} \ldots$ doesnt converge to x^*

This means $\exists$ a subsequence x^{k_n} that converges to some $\bar{x} \neq x^*$

Since $\bar{x} \neq x^*$ and f is convex

$f(\bar{x}) > f(x^*) \to ①$

Let's use the same taylor theorem with lagrange reminder

$f(x) = f(\bar{x}) + f'(c)(x - x^*)$ ($\because$ c is a number b/w x and x^*

$f'(x^k)$ converges to zero

Choose large k_n s.t. $f'(c) = 0 \; \forall \; x^{k_n}$

using taylor theorem

$f(x^{(k_n)}) = f(x^*) + f'(c) \cdot (x^{k_n} - x^*)$

now, $\lim_{n \to \infty} f(x^{k_n}) \to f'(x)$ using continuity of f

hence $f(\bar{x}) = f(x^*)$ so it is contradiction

$\therefore$ Thus assumption $x^{(0)}, x^{(1)}, x^{(2)}$ doesn't converge to x^* must be false & conclude that

~~x_0, x_1, x_2~~ $x^{(0)}, x^{(1)}, x^{(2)}$ converge to x^*

6) b) Yes, since f is a strong convex function, it implies that local maxima is the global maxima

Now,

$f(x^{(0)}) > f(x^{(1)}) \cdots > f(x^{(k)})$ we need to prove that $x^k \to x^*$. Let $x^k \to y^\circ$ where $y \neq x^*$ then $f'(x^{(k)})$ doesn't converge from (9)

Now $x^{(k+1)} = x^k + cf'(x^k)$

Taking limit $k \to \infty$

$$\lim_{k\to\infty} x^{k+1} = \lim_{k\to\infty} x^k + \lim_{k\to\infty} cf'(x^k)$$

$$y = y + cf'(y)$$

$$0 = cf'(y) \Rightarrow f'(y) = 0$$

But $f'(x^k)$ shouldn't converge to 0. hence a contradiction

hence $x^{(k)} \to x^*$

Q3) we are given that $A \in \mathbb{R}^{m \times m}$ $\rightarrow$ non symmetric matrix

also

$A = QTQ^T$ and all $t_{11}, t_{22} \dots t_{mm}$ are distinct.

Since T is a upper ~~[illegible]~~ Δ matrix $t_{11}, t_{22} \dots t_{mm}$ are the distinct eigenvalues of T and since A and T are similar A will also have the same eigenvalues.

* Also $A = QTQ^T \Rightarrow T = Q^T A Q$

Now let v be an eigenvector corresponding to eigenvalue λ of T then, $Tv = \lambda v$. we claim that Qv will be an eigenvector of A with same eigenvalue λ.

Proof :- $A(Qv) = QTQ^TQv = QTv = \lambda Qv$

$$\therefore \; A(Qv) = \lambda(Qv)$$

Hence if we find an eigenvector corresponding to T, we will find the eigenvector corresponding to A.

In order to find the eigenvector of T corresponding to $\lambda = t_{ii}$, we will apply the inverse power iteration method with $\mu \approx t_{ii}$ since it satisfies the condition of power iteration method

Since all eigenvalues are unique

$$\left|\frac{1}{t_{ii} - \mu_i}\right| >> \left|\frac{1}{t_{jj} - \mu_i}\right| \quad \forall \; j \neq i$$

The steps of the inverse power iteration method are

① Start with initial guess $x \to 0$

② Compute the product

$$(A - \mu I)\vec{x}_{k+1} = \vec{x}_k$$

③ Then apply normalize $\vec{x}_{k+1} = \frac{\vec{x}_{k+1}}{\|\vec{x}_{k+1}\|}$

④ repeat step 2 and 3 until convergence

The resulting $\vec{x}$ will point towards the eigenvector e_i with eigenvalue t_{ii} since $\left|\frac{1}{t_{ii}-\mu_i}\right|$ is the highest eigenvalue and Power iteration points towards highest eigenvalue eigenvector.

The convergence rate will be

$$\cancel{\max_{j \neq i} \frac{t_{ij} - \mu}{t_{jj} - \mu}} \qquad \max_{j \neq i} \frac{|t_{ij} - \mu|}{|t_{jj} - \mu|}$$

The total no. of operations required will be $O(m)$ per iteration. Since T is an upper $\triangle$ matrix and we know that step 2 & 3 takes $O(m)$ steps.

The total no of iterations depend on convergence rate but its $O(m)$

Inorder to calculate the required eigenvector we will find $\vec{y} = Q\vec{x}$ (obtained from inverse power iteration)

$\hookrightarrow$ it takes $O(m^2)$ flops

hence total operations is $O(m^2)$

5) a) To avoid Runge's phenomenon, we can choose sample points as zeros of the Chebyshev polynomial of first kind.

The Chebyshev nodes provide an optimal distribution of points that minimize the interpolation error.

In [-1,1] the nodes are given by (Chebyshev)

$$x_k) = \cos\left(\left(2k-1\right)\pi/_{2n}\right) \quad k = 1, 2, \ldots, n$$

given that 3 sample points has to be taken, (n=3)

$$x_1 = \cos\left(2\times1-1\right)\pi/_{2\times3}\Big) = \cos\left(\pi/_6\right) = \sqrt{3}/_2$$

$$x_2 = \cos\left(2\times2-1\right)\pi/_{2\times3}\Big) = \cos\left(\pi/_2\right) = 0$$

$$x_3 = \cos\left(\left(2\times3-1\right)\pi/_{2\times3}\right) = \cos\left(5\pi/_6\right) = -\sqrt{3}/_2$$

∴ The coordinates are

$$\left(\sqrt{3}/_2, 0, -\sqrt{3}/_2\right)$$

b) Now, construct a system of equations the quadrature rule ~~&~~ for polynomial of increasing degree.

Let w_i be weights and points as x_i

For degree = 0, (constant function)

$$w_{1}.1 + w_{2}.1 + w_{3}.1 = 2$$

For degree = 1 (linear functions)

$$w_1.x_1 + w_2x_2 + w_3x_3 = 0$$

For degree = 2 (quadratic fn)

$$w_1x_1^2 + w_2x_2^2 + w_3x_3^2 = 2/3$$

For degree = 3 (cubic fn)

$$w_1x_1^3 + w_2x_2^3 + w_3x_3^3 = 0$$

For (degree = 4) $w_1x_1^4 + w_2x_2^4 + w_3x_3^4 = 2/5$

for degree =5 (quintic fns)

$\omega_1 x_1^5 + \omega_2 x_2^5 + \omega_3 x_3^5 = 0$

by putting the x_1, x_2 and x_3 as in part 3.
we get the highest degree as 3

$\omega_1 = 4/9$, $\omega_2 = 10/9$, $\omega_3 = 4/9$

2) b) The algorithm

Step ①:- compute the SVD of $A = U\Sigma A^*$
this is done using any standard SVD algorithm

Step ②: compute $c = U^* b$
simple matrix vector multiplication

Step ③:- solve $\Sigma d = c$ for d
This is done by solving the system of linear eqs.
Since Σ is diagonal (easy to do: just divide each element of 'C' by corresponding element of Σ)

Step ④:- compute $y = Vd$ (simple matrix vector mult).

The above algorithm has time complexity of $O(mnr)$ where 'm' is the no. of rows in A
n = no. of columns in A & 'r' is rank of A.

2 a)

1) a)

The columns of Q and R in a reduced QR factorisation of $A = QR$ are ill conditioned w.r.to perturbations in corresponding columns of 'A' if A matrix is ill conditioned.

4) The fixed points of iteration values of (x, z, u) that satisfies $x = \left(\frac{1}{2}\right) * (z+4)$, $\rightarrow$ ①

~~$z = \max\{(x+4)/2, 9\}$~~ ②

$z = \max\{(x+4)/2, a\} \rightarrow$ ②

$u = u + x + z \rightarrow$ ③

subs ① into ③.

$u = 2x + 24$.

Rearrange this eq., $x = -4$ (put it into ① eq)

$z = -4$

by substituting in individual second eq gives

$\max\{(-4/2, 9) = -4$

$\therefore$ if $a \leq 0$

then fixed point = $(0,0,0)$

if $a > 0$, fixed point = $(a/2, a, -3a/2)$

analyse the convergence rate

$$x^k = (x^k, z^k, u^k)$$

iteration converges to $x^* = [x^*, z^*, u^*]$

$$x^{k+1} - x^* = f(x^k) - f(x^*) \approx J(x^*)(x^k - x^*)$$

so, Jacobian matrix of iteration is

$$J(x) = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

eigen values of $J(x^*)$ $= -1, \ (1 \pm \sqrt{5})/4$

$\therefore \ a > (3 + \sqrt{5}/2)$ or $a < (3 - \sqrt{5}/2)$