Math 128B, K Miller, Spr 2001
Final Exam, May 16 Name Keith Miller

| | |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 | |
| 6 | |
| total | |

1a° Give the defn of the matrix operator norm $\|A\|_\infty$ for an $n \times n$ matrix $A$. For the particular $A$ below, show that $\|A\|_\infty = 6$ and find a vector $x \neq 0$ such that $\|Ax\|_\infty = 6\|x\|_\infty$.

$$A = \begin{pmatrix} -1 & 2 & 1 \\ 2 & -3 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

Answer. Defn $\|A\|_\infty \equiv \sup_{x \neq 0} \dfrac{\|Ax\|_\infty}{\|x\|_\infty}$.

- $$Ax_1 = \begin{pmatrix} -1x_1 + 2x_2 + 1x_3 \\ 2x_1 - 3x_2 - 1x_3 \\ 1x_1 - 1x_2 + 2x_3 \end{pmatrix} \begin{array}{l} \leftarrow |\ | \le (1+2+1)\|x\|_\infty \\ \leftarrow |\ | \le (2+3+1)\|x\|_\infty \\ \leftarrow |\ | \le (1+1+2)\|x\|_\infty \end{array}$$

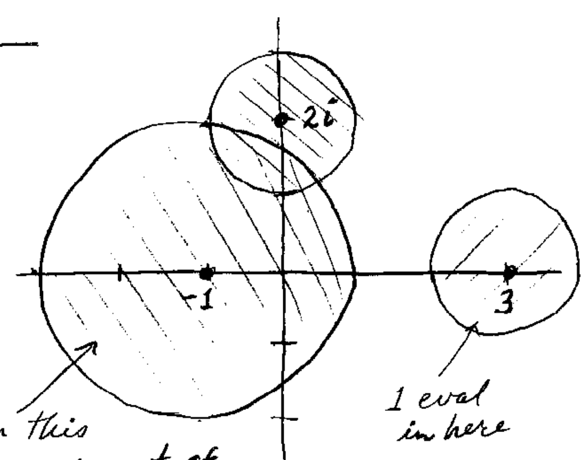
Thus $\|Ax\|_\infty \le 6\|x\|_\infty$ for all $x$, or ① $\dfrac{\|Ax\|_\infty}{\|x\|_\infty} \le 6$ for all $x \neq 0$.

- However equality can occur in ① for the vector $x = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ which has $\|x\|_\infty = 1$ and $\|Ax\|_\infty = 6$.

Thus $6 = \max_{x \neq 0} \dfrac{\|Ax\|_\infty}{\|x\|_\infty}$

1b° Consider $A = \begin{pmatrix} 3 & i & 0 \\ 0 & 2i & -1 \\ 1 & 1 & -1 \end{pmatrix}$ Draw in the Gershgorin row circles on the complex plane shown. Then shade in the region where the eigenvalues of $A$ can be found. Tell how many eigenvalues (algebraic multiplicities considered) are in each part.

Answer The row circles about $3, 2i, -1$ as centers have radii $1, 1, 2$.




2 evals in this connected component of the union of the row circles

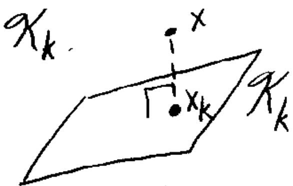
1 eval in here

Math 128B Final, K-M, Sp. 01 Name ____________________

2° We consider the Conjugate Gradient method for solving $Ax = r_0$, where $A$ is $n \times n$, symmetric, pos def. On $R^n$ we use the "energy" inner product and norm $((u,v)) \equiv (Au, v)$, $|||u||| \equiv ((u,u))^{1/2}$. We say $u$ and $v$ are "conjugate-orthogonal" if $((u,v)) = 0$.

At the end of the kth step of CG:

- We have previously computed the conj-orthog vectors $p_1, \cdots p_k$ such that $\text{span}\{p_1, \cdots, p_k\} = \text{span}\{r_0, Ar_0, \cdots A^{k-1} r_0\} \equiv \mathcal{K}_k$.
- We have computed $x_k$ to be that elt of $\mathcal{K}_k$ which best approximates the unknown $x$ in energy norm.




- We have computed its residual $r_k \equiv r_0 - Ax_k$ and find that $r_k \neq 0$.
- We have saved $r_0$, $p_k$, $(Ap_k, p_k)$, $x_k$, $r_k$ (but have discarded the other previous vectors.

Now, for the (k+1)st step of CG:

2a° Why is $r_k$ regular-orthogonal to $\mathcal{K}_k$? (Hint: $x - x_k$ is conj-orthog to $\mathcal{K}_k$, as indicated in the diagram above.)

2b° Why is $\{p_1, \cdots p_k, r_k\}$ a basis for $\mathcal{K}_{k+1}$?

2c° Explain why $r_k$ is already conj-orthog to some of the previous $p_j$'s.

2d° Thus, explain how one computes $p_{k+1}$ so as to be conj-orthog to $p_1, \cdots p_k$.

2e° Then, how does one efficiently compute $x_{k+1}$ and $r_{k+1}$? (Note efficiency requires that within each CG step we apply A to only one vector)

Answers

2a The orthogonality condition for $x_k$ to be the best approx in $\mathcal{K}_k$ to $x$ in the energy norm is that

$$((x - x_k, v)) = 0 \quad \text{for all } v \text{ in } \mathcal{K}_k$$

$$\equiv (\underbrace{A(x - x_k)}_{= r_0 - Ax_k = r_k}, v)$$

Thus $r_k$ is regular-orthogonal to $\mathcal{K}_k$

2b First, $r_k = (r_0 - A \underbrace{x_k}_{\in \mathcal{K}_k})$ is in $\mathcal{K}_{k+1}$.

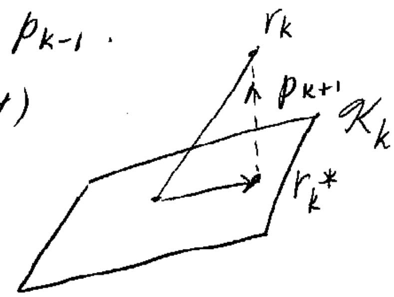
Second $r_k$ is regular-orthog to the $p_1, \dots p_k$, so is not a lin combination of them. Hence $\{p_1, \dots p_k, r_k\}$ are $k+1$ lin ind vectors in a space $\mathcal{K}_{k+1}$ of dimension $\leq k+1$, so are a basis for that space.

2c $A(\mathcal{K}_{k-1}) \subset \mathcal{K}_k$. Thus $Ap_j \in \mathcal{K}_k$ for $j = 1, \dots k-1$.

But $r_k$ is reg-orthog to $\mathcal{K}_k$, hence

$$\underbrace{(r_k, Ap_j)}_{= (Ar_k, p_j) \equiv ((r_k, p_j))} = 0 \quad \text{for } j = 1, \dots k-1$$

Hence $r_k$ is already conj-orthog to $p_1, \dots p_{k-1}$.




2d Gram-Schmidt (in the (( , )) inner product) would require us to orthogonalize $r_k$ against all the previous vectors $p_1 \dots p_k$:

I.e. $p_{k+1} = r_k - r_k^*$ where

$r_k^*$ is the conj-projection of $r_k$ onto span$\{p_1 \dots p_k\}$

But since $r_k$ is conj orthogonal to $p_1 \dots p_{k-1}$, this projection has only the single component.

$$r_k^* = \beta_{k+1} p_k \quad \text{where (1) } \beta_{k+1} = \frac{((r_k, p_k))}{((p_k, p_k))} = \frac{(r_k, Ap_k)}{(Ap_k, p_k)}$$

Then.

$$\text{(2) } p_{k+1} = r_k - \beta_{k+1} p_k$$

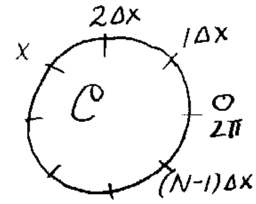
2e Now one computes (3) $Ap_{k+1}$ (the only "application of A")

Then (5) $x_{k+1} = x_k + \alpha_{k+1} p_{k+1}$,

where (4) $$\alpha_{k+1} = \frac{((x, p_{k+1}))}{((p_{k+1}, p_{k+1}))} = \frac{(\overbrace{Ax}^{r_0}, p_{k+1})}{(Ap_{k+1}, p_{k+1})} = \frac{(r_0, p_{k+1})}{(Ap_{k+1}, p_{k+1})}$$

Then (6) $$r_{k+1} = r_0 - A \underbrace{x_{k+1}}_{x_k + \alpha_{k+1} p_{k+1}} = \underbrace{(r_0 - Ax_k)}_{\equiv r_k} - \alpha_{k+1} A p_{k+1}.$$

Math 128B Final, "KM", Spr 01 Name Keith Miller




3° Let $f$ be a discrete complex-valued fn on the discrete $[0, 2\pi)$ circle $C$ with $N$ nodes and $\Delta x = 2\pi/N$, where $N = 2^k$ is a power of 2. The FFT involves a fast way to compute the $N$ Fourier coefficients

① $c_j \equiv \sum_{x \in C} f(x) e^{-ijx}$ , $j = 0, 1, \ldots N-1$.

3a° The basis of FFT is the identity

② $c_j = \tilde{c}_{j1} + w^j \hat{c}_{j1}$ .

What are $w$, $j1$, $\tilde{c}_{j1}$, $\hat{c}_{j1}$ ?

3b° Let $p_k$ denote the number of complex multiplications required to do a FFT on $N = 2^k$ nodes. Using ② find a formula (or at least a good upper bound) for $p_k$.

---

Answer 3a° $w \equiv e^{-i\Delta x}$
$j1$ = the remainder of $j$ after division by $N/2$,
$j1 = 0, 1, 2, \ldots N/2 - 1$.
$\tilde{c}_{j1}$ and $\hat{c}_{j1}$ are the Fourier coeffs of $\tilde{f}$ and $\hat{f}$ (the restrictions of $f$ to the even and the odd nodes of $C$).

Answer 3b° Because of ②, the computation of a FFT on $N = 2^k$ nodes requires two computations of FFTs on $N/2 = 2^{k-1}$ nodes (for $\tilde{f}$ and $\hat{f}$ respectively) plus the $N = 2^k$ mults by $w^j$ in ② to combine them. Thus

③ $p_k = 2 \cdot p_k + 2^k$

Now $p_0 = 0$ (because for FT on $2^0 = 1$ node the single coeff $c_0$ is just the value of $f$ itself at that node)

by ③ $p_1 = 2 \cdot p_0 + 2^1 = 2 \cdot 0 + 2^1 = 1 \cdot 2^1$

by ③ $p_2 = 2 \cdot p_1 + 2^2 = 2(1 \cdot 2^1) + 2^2 = 2 \cdot 2^2$

by ③ $p_3 = 2 \cdot p_2 + 2^3 = 2 \cdot 2 \, 2^2 + 2^3 = (2+1) 2^3 = 3 \cdot 2^3$

by ③ $p_4 = 2 \cdot p_3 + 2^4 = 2(3 \cdot 2^3) + 2^4 = (3+1) 2^4 = 4 \cdot 2^4$

$\vdots$

by ③ $\underline{p_k} = 2 \cdot p_{k-1} + 2^k = 2((k-1) \cdot 2^{k-1} + 2^k = ((k-1)+1) 2^k = \underline{k \cdot 2^k}$

$= \underline{(\log_2 N) \cdot N}$

Math 128B Final, K M Spr 01 Name Keith Miller

4a° A Householder reflection matrix is of the form $H = I - 2vv^T$ where $v$ is a unit vector in $R^n$. Show that $H$ is symmetric ($H^T = H$), self-inverse ($HH = I$), and thus orthogonal ($H^TH = I$).

4b° Applying such reflections on the left to col vectors is quick. Efficiently compute $Hx$, where $x = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 2 \end{pmatrix}$ and $v = \frac{w}{\|w\|}$ with $w = \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 2 \end{pmatrix}$.

(Note $vv^T = \frac{1}{\|w\|^2} ww^T$, so you never need square roots.

4c° Likewise, applying reflections on the right to row vectors is quick. Find $x^TH$ where $x$ and $H$ are as above in (4b).

4d° Such reflections can be used to zero out all but one component of a vector. Find the unit vector $v = \frac{1}{\|w\|} w$ for the reflection $H$ which maps $d = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$ into $Hd = \begin{pmatrix} * \\ 0 \\ 0 \end{pmatrix}$, using the proper $\pm$ convention to avoid roundoff error.

---

Answers

4a $H^T = (I - 2vv^T)^T = I^T - 2(vv^T)^T = I - 2(\underbrace{v^{TT}}_{=v} v^T) = H$

$$HH = (I - 2vv^T)(I - 2vv^T) = I - 4(vv^T) + 4\underbrace{(vv^T)(vv^T)}_{\underbrace{v}_{n\times 1}\underbrace{(v^Tv)}_{=1\ (1\times 1 \text{ matrix})}\underbrace{v^T}_{1\times n}} = I$$

$H^TH = HH = I$

4b The point of this part is that application of $H$ to a vector $x$ is quick because it requires only computing a single dot product ($O(n)$ ops) since $Hx = x - 2v\underbrace{(v^Tx)}_{\text{this is just } v\cdot x} = x$ - a multiple of $v$.

Here I goofed in recopying the problem + gave you an $x$ in the direction of $w$ (in fact $x = w$) so of course its reflection in $w^\perp$ is just $-x$. I meant to give you $x = \begin{pmatrix} 2 \\ -2 \\ 2 \\ 1 \\ 1 \end{pmatrix}$ for which $\|w\|^2 = 10$, $w \cdot x = 10$. Thus for this $x$, same $w$,

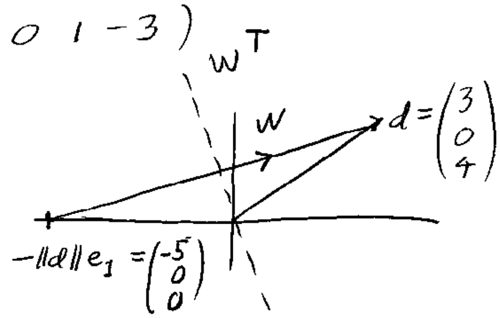
$$Hx = x - \underbrace{\frac{2}{\|w\|^2}}_{10} w \underbrace{(w^Tx)}_{10} = x - 2w = \begin{pmatrix} 2 \\ -2 \\ 2 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ -4 \\ 2 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \\ -3 \end{pmatrix}$$

4c The point of this part is that to apply $H$ on the right to a row vector involves merely converting the row to a column, applying $H$ on left as in 4b, then converting back to a row.

$$(Hx)^T = (x^T \underbrace{H^T}_{=I}) = \text{the desired } (x^T H)$$

Thus here $(x^T H) = (Hx)^T = (0\ 2\ 0\ 1\ -3)$

4d We map $d = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$ into $\begin{pmatrix} -\|d\| \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}$




Thus $w = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} - \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ 4 \end{pmatrix}$

and $\|w\|^2 = 80$, $v = \frac{1}{\sqrt{80}} w$, and

$$H = \left(I - \frac{2}{80} w w^T\right)$$

Math 128B Final, KM, Spr 01 Name ______________________

<u>5a°</u> Let $A_1$ be a real $4\times4$ full matrix. <u>Explain</u> (referring to (4d)) how application of 3 Householder reflections, first $H_1$, then $H_2$, then $H_3$ allows one to compute the $Q_1R_1$ decomposition of $A_1$.
Write $R_1$ in terms of $A_1, H_1, H_2, H_3$, using <u>parentheses</u> to indicate the order of application of the $H$'s.
Then write $Q_1$ in terms of the $H$'s. <u>Why</u> is $Q_1$ orthogonal?

<u>5b°</u> Use $R_1$ and these saved $H$'s (ie save their $v$ vectors) to transform $A_1$ into an orthogonally similar matrix $A_2$ in the first stage of "the QR algorithm for finding eigenvalues".
<u>Show</u> that $A_2$ <u>is</u> orthogonally similar to $A_1$. Use <u>parentheses</u> again to indicate the order of applying the $H$'s.

<u>5c°</u> Refer to (4b) and (4c) to <u>explain</u> why "applying the $H$'s to matrices" in (5a) and (5b) is efficient

---

<u>Answer 5a</u>

$$A_1 = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \qquad H_1A_1 = \begin{pmatrix} * & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix}$$

$H_1$ maps this into $\begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix}$ as in (4d).

$H_2$ maps this 3-vector into $\begin{pmatrix} * \\ 0 \\ 0 \end{pmatrix}$

$$H_2(H_1A) = \begin{pmatrix} * & x & x & x \\ 0 & * & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{pmatrix}$$

$H_3$ maps this 2 vector into $\begin{pmatrix} * \\ 0 \end{pmatrix}$

$$① \; H_3\left(H_2\left(H_1(A_1)\right)\right) = R_1 = \begin{pmatrix} * & x & x & x \\ & * & x & x \\ 0 & & * & x \\ & & & x \end{pmatrix}$$

Now the $H$'s are self-inverse, so applying $H_1H_2H_3$ to ① we get

$$A_1 = \cancel{H_1}\cancel{H_2}\cancel{H_3}\cancel{H_3}\cancel{H_2}\cancel{H_1}A_1 = \underbrace{H_1H_2H_3}\, R_1$$

② <u>this</u> is our $Q_1$. Note that <u>the product</u> of orthog matrices is orthogonal

<u>Note</u> ③ $A_1 = Q_1R_1$, hence $Q_1^TA_1 = R_1$ since $Q_1^TQ_1 = I$

Keith Miller

Answer 5b $A_2$ is defined by

(4) $A_2 \equiv R_1 Q_1$ , which $\overset{\text{by (3)}}{=} Q_1^T A Q_1$ ,

so $A_2$ is orthogonally similar to $A_1$.

We compute by applying the H's on the left

$$A_2 \equiv ((R_1)\underbrace{H_1) H_2) H_3}_{Q_1}$$

Answer 5c Applying an H on the left in (5a) to a matrix "B"

applying H to each col of B, as we saw in (4b) that

requires only $\approx n$ ops for each col vector.

Thus to apply the $H_1, H_2, \cdots H_{n-1}$ on the left in (5a)

requires $\approx n^2 + (n-1)^2 + \cdots (2)^2 \approx \frac{1}{3}n^3$ ops.

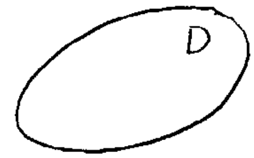
Applying an H on the right in (5b) as we saw

in (4b) is exactly the same except we apply H to the row

vectors of B. Hence also $\approx \frac{1}{3}n^3$ ops.

Math 128B Final, "K"M, Spr 01 Name Keith Miller

6 Consider the Dirichlet problem for Poisson's eqn




$$① \begin{cases} -\Delta u = f(x) \text{ in } D, \\ u = 0 \quad \text{on } \partial D. \end{cases}$$

6a What is the function space $pw\,C^1_z(\bar{D})$?
Give the defn of " $u$ is a weak soln of ① ".

6b Let $M_h$ be a finite dimensional subspace of $pw\,C^1_z(\bar{D})$. Define the Galerkin or Finite Element approximate soln $U$.

6c State and prove the " best approximation property in energy norm" for the FE soln $U$. ( But first define the "energy inner product and norm $(( \ , \ ))$ and $||| \ \ |||$.")

---

Answer 6a

Defn The space $pw\,C^1_z(\bar{D})$ is the space of continuous piecewise $C^1$ fns on $\bar{D}$ with zero bndry values.

Defn $u$ is a weak soln of ① if

$$② \begin{cases} u \in pw\,C^1_z(\bar{D}) \\ a(u,\varphi) = (f,\varphi) \text{ for all test fns } \varphi \text{ in } pw\,C^1_z(\bar{D}). \end{cases}$$

Defn for all $v, w$ in $pw\,C^{01}_z$ we define

$$a(v,w) = \int_D \nabla v \cdot \nabla w \, dx, \quad ((v,w)) \equiv a(v,w), \quad |||v||| \equiv ((v,v))^{\frac{1}{2}}.$$

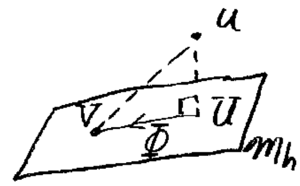
Answer 6b Defn The FE soln $U$ is the unique soln of

$$③ \begin{cases} U \in M_h \\ a(U,\Phi) = (f,\Phi) \text{ for all test fns } \varphi \text{ in } M_h. \end{cases}$$

Answer 6c

- Since (( , )) is an inner product on $pw\,C^1_Z(\bar{D})$ we have the usual Cauchy-Schwarz inequality $((v,w)) \le |||v|||\;|||w|||$
- Since we can use test fns $\Phi$ in (3) also, subtracting (4) from (3) we get

(5) $((u-U,\Phi)) \equiv a(u-U,\Phi) = 0$ for all $\Phi$ in $M_h$.




Thm (Best approx property) For any elt V in $M_h$ we have $|||u-U||| \le |||u-V|||$

Proof $\Phi \equiv V-U$ is in $M_h$. Hence.

$$\|u-U\|^2 = ((u-U, \underbrace{u-U}_{u-V+\Phi})) = ((u-U, u-V)) + \underbrace{((u-U,\Phi))}_{=0 \text{ by } (5)}$$

$$\underset{\text{Schwarz Inequality}}{\overset{\text{Cauchy}}{\le}} |||u-U|||\;|||u-V|||.$$

Better Proof Use Pythagoras and the orthogonality of $u-U$ and $\phi$ in the energy inner product, as shown in the diagram.

$$|||u-V|||^2 = |||(u-U)+\Phi|||^2 \equiv (((u-u)+\Phi, (u-U)+\Phi))$$

$$= |||u-u|||^2 + 2\underbrace{((u-U,\Phi))}_{=0 \text{ by } (5)} + \underbrace{|||\Phi|||^2}_{>0 \text{ if } \Phi \neq 0}$$

$$> |||u-U|||^2.$$

qed

END