Nicholas Miller

MIS HW #5

#10.15 Let A be an $m\times n$ matrix with columns $a_1, a_2, \dots, a_n$

Then $G = A^TA$ where $G_{ij} = a_i^T a_j$

$$G^T = (A^TA)^T = A^T(A^T)^T = A^TA = G$$

Then $G_{ij} = G_{ji}^T = G_{ji}$

Let $\|a_i - a_j\|^2 = (a_i - a_j)^T(a_i - a_j)$

$$= (a_i^T - a_j^T)(a_i - a_j)$$

$$= a_i^Ta_i - 2a_i^Ta_j + a_j^Ta_j$$

$$= G_{ii} - 2G_{ij} + G_{jj}$$

Taking the square root of both sides

$$\|a_i - a_j\| = \sqrt{G_{ii} - 2G_{ij} + G_{jj}}$$

#10.16 (a) $A = [a_1 \dots a_k]$

$$\mu = [\mu_{a_1} \dots \mu_{a_k}] = \left[\frac{\mathbf{1}^Ta_1}{n} \dots \frac{\mathbf{1}^Ta_k}{n}\right]$$

(b) $\tilde{A} = [a_1 - \mathbf{1}^T\mu_1 \dots a_k - \mathbf{1}^T\mu_k]$

(c) case 1: $i=j$

$$\Sigma = \frac{1}{N}\tilde{a}_i^T\tilde{a}_i = \frac{\tilde{a}_i^T}{\sigma_{a_i}\sqrt{N}} \times \frac{\tilde{a}_i}{\sigma_{a_i}\sqrt{N}} \times \sigma_{a_i}^2 = (z_{a_i} \times z_{a_i}) \times \mathrm{std}\, a_i^2 = \mathrm{std}(a_i)^2$$

case 2: $i \neq j$

$$\Sigma = \frac{1}{N}\tilde{a}_i^T\tilde{a}_j = \frac{\tilde{a}_i^T}{\sigma_{a_i}\sqrt{N}} \times \frac{\tilde{a}_j^T}{\sigma_{a_j}\sqrt{N}} \times \sigma_{a_i} \times \sigma_{a_j} = \rho_{ij} \times \mathrm{std}(a_i) \times \mathrm{std}(a_j)$$

(d)

$$Z = \left[\frac{a_1 - \mathbb{1}^T\mu_1}{std(a_1)} \; \cdots \; \frac{a_K - \mathbb{1}^T\mu_K}{std(a_K)}\right]$$

#10.36 (a) let $Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$Ax = \begin{pmatrix} \sum_{i=1}^{n} a_{1i}x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi}x_i \end{pmatrix}$$

$$x^TAx = (x_1 \cdots x_n)\begin{pmatrix} \sum a_{1i}x_i \\ \vdots \\ \sum a_{mi}x_i \end{pmatrix}$$

$$= \sum a_{1i}x_ix_1 + \cdots + \sum a_{mi}x_ix_n$$

$$= \sum_{i,j=1}^{n} a_{ji}x_ix_j$$

(b) $x^TA^Tx = (Ax)^Tx = \langle Ax, x\rangle = \langle x, Ax\rangle = x^TAx$

(c) $x^T\left(\frac{A+A^T}{2}\right)x = \frac{x^TAx}{2} + \frac{x^TA^Tx}{2} = \frac{x^TAx}{2} + \frac{x^TAx}{2} = x^TAx$

(d) $2x_1^2 - 3x_1x_2 - x_2^2 = 2x_1^2 - \frac{3}{2}x_1x_2 - \frac{3}{2}x_2x_1 - x_2^2$

$$= (x_1\; x_2)\begin{pmatrix} 2 & -3/2 \\ -3/2 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x^TAx$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ & $A = \begin{pmatrix} 2 & -3/2 \\ -3/2 & -1 \end{pmatrix}$

Scanned with CamScanner

#10.42 let $A = I + ab^T$

Then $Ax = (I + ab^T)x$
$= Ix + ab^Tx$
① ②

① $Ix = x$ I is very sparse so this will Take $nnz(I) = n$ flops

② b^Tx: takes n mult. & n-1 add = 2n-1 flops
$a(c)$: takes n mult & n-1 add = 2n-1 flops

Sum through we get n more flops
so the whole process takes
$n + 2(2n-1) + n$ flops which
is $O(n)$

So if we distribute x and carry out calculations this will be done in linear time.

#11.3 (a) $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $Y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$AX = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $AY = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
$AX = AY$ but $X \neq Y$

(b) $AX = AY \Leftrightarrow A^{-1}AX = A^{-1}AY \Leftrightarrow X = Y$

(c) example in (a) let $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is A invertible

$\det(A) = 0$ so A is not invertible then not left invertible

#11.9 (a) $(I+BA)x = 0 \Leftrightarrow x + B(Ax) = x$

$\Leftrightarrow x + By = 0 \Leftrightarrow Ax + ABy = 0 \Leftrightarrow y + ABy = 0$

$\Leftrightarrow (I+AB)y = 0$

But $I+AB$ is invertible thus $y = Ax = 0$

Now $x + B(Ax) = 0 \Leftrightarrow x + B(0) = 0 \Leftrightarrow x = 0$

Hence the null space of $I+BA$ is $\{0\}$ and thus $I+BA$ is invertible

(b) $B(I+AB) = B + BAB - (I+BA)B$

$$B(I+AB) = (I+BA)B$$
$$B(I+AB)(I+AB)^{-1} = (I+BA)B(I+AB^{-1})$$
$$B = (I+BA)B(I+AB)^{-1}$$
$$(I+BA)^{-1}B = (I+BA)^{-1}(I+BA)B(I+AB)^{-1}$$
$$(I+BA)^{-1}B = B(I+AB)^{-1}$$

#11.27 $n = 500$

time 0.372 s magnitude $= 3.24 \times 10^{-11}$

$$\frac{2 \cdot 500^3 \text{ flops}}{0.372 s} \cdot \frac{\text{Gflops}}{10^9 \text{ flops}} = 0.672 \text{ Gflops/s}$$

$n = 1000$

time 0.603 s magnitude $= 3.24 \times 10^{-11}$

$$\frac{2 \cdot 1000^3 \text{ flops}}{0.603 s} \cdot \frac{\text{Gflops}}{10^9 \text{ flops}} = 3.317 \text{ Gflops/s}$$

$n = 2000$ time 0.890 s magnitude 1.42×10^{-11}

$$\frac{2 \cdot 2000^3 \text{ flops}}{0.890 s} \cdot \frac{\text{Gflops}}{10^9 \text{ flops}} = 17.977 \text{ Gflops/s}$$

Scanned with CamScanner

#12.13 (a) Let $x^{(k+1)} = x^{(k)}$

Then $x^{(k+1)} = x^{(k)} - \mu A^T (Ax^{(k)} - b)$

$$\Downarrow$$

$$x^{(k)} = x^{(k)} - \mu A^T (Ax^{(k)} - b)$$

$$\Downarrow$$

$$0 = -\mu A^T (Ax^{(k)} - b)$$

$$\Downarrow$$

$$0 = A^T (Ax^{(k)} - b)$$

Thus $(Ax^{(k)} - b)$ is minimized

So $x^{(k)} = \hat{x}$

(b) $x^{(k+1)} = x^{(k)} - \mu A^T (Ax^{(k)} - b)$

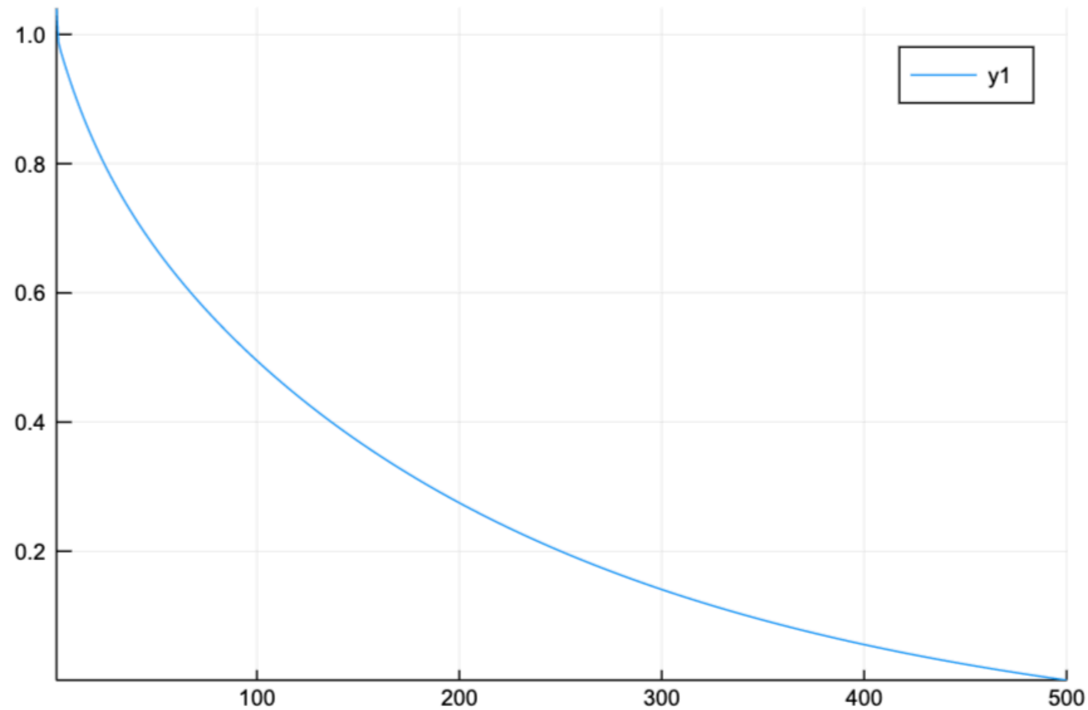
$= x^{(k)} - \mu A^T A x^{(k)} - \mu A^T b$

$= (I - \mu A^T A) x^{(k)} - \mu A^T b$

$\Rightarrow x^{(k+1)} = F x^{(k)} + g$

where $F = (I - \mu A^T A)$ & $g = -\mu A^T b$

(c)



As we can see after each iteration Ax-b is converging to zero.