CS205 Review Session #1 Notes

1 HW 1.1 Hints

1.1 Addition Example

Recall that the relative error is defined as:

$$\text{relative error} = \left|\frac{\text{computed\_ result} - \text{analytic\_ result}}{\text{analytic\_ result}}\right|$$

Consider the relative error induced by the addition of two positive numbers x and y. The simple act of storing each number in floating-point introduces some error yielding $\bar{x} = (1+\epsilon_1)x$ and $\bar{y} = (1+\epsilon_2)y$. Performing the addition induces some roundoff error both for the operation itself as well as the storage of the result, which we will model as $(1+\epsilon_3)(\bar{x}+\bar{y})$. So:

$$\begin{aligned}
R.E. &= \left|\frac{(1+\epsilon_3)((1+\epsilon_1)x + (1+\epsilon_2)y) - (x+y)}{x+y}\right| \\
&\leq \left|\frac{(1+\epsilon_3)(1+\epsilon_4)(x+y) - (x+y)}{x+y}\right| \\
&= |(1+\epsilon_3)(1+\epsilon_4) - 1| \\
&= |(1+\epsilon_5)^2 - 1| \\
&\leq (1+\epsilon_{\text{max}})^2 - 1 \\
&= 2\epsilon_{\text{max}} + O(\epsilon_{\text{max}}^2)
\end{aligned}$$

1.2 Aside

Consider a sequence of multiplied rounding error factors:

$$(1+\epsilon_1)(1+\epsilon_2)\cdots(1+\epsilon_k)$$

It's clear that the cumulative error can be bounded both above and below by:

$$(1-\epsilon_{\text{max}})^k \leq (1+\epsilon_1)(1+\epsilon_2)\cdots(1+\epsilon_k) \leq (1+\epsilon_{\text{max}})^k$$

From this bound, we can see that there must exist some ϵ_* that satisfies:

$$(1+\epsilon_*)^k = (1+\epsilon_1)\cdots(1+\epsilon_k)$$

1.3 Multiplication Series Example

Suppose we wish to calculate the relative error introduced in the computation of x^n given the recurrence:

$$\begin{aligned} m_1 &= x \\ m_k &= m_{k-1}x \end{aligned}$$

We know that x will be stored on the computer as $\bar{x} = (1+\epsilon)x$. Thus, we can write the evaluation as another recurrence

$$\begin{aligned} s_1 &= (1+\epsilon)x \\ s_k &= [s_{k-1}(1+\epsilon)x](1+\epsilon_k) \end{aligned}$$

whose solution is clearly

$$s_n = x^n(1+\epsilon)^n(1+\epsilon_2)(1+\epsilon_3)\cdots(1+\epsilon_n)$$

which simplifies (for some ϵ_*) to:

$$s_n = x^n(1+\epsilon_*)^{2n-1} = x^n(1+(2n-1)\epsilon_* + O(\epsilon_*^2))$$

So, the relative error is:

$$RE = \left|[x^n(1+(2n-1)\epsilon_* + O(\epsilon_*^2)) - x^n]/x^n\right| \leq (2n-1)\epsilon_{\text{max}} + O(\epsilon_{\text{max}}^2)$$

2 HW 1.2 Hints

Recall that permutation matrices are identity matrices that have had their rows (or columns) interchanged. Clearly a permutation matrix permuted by itself will yield the identity.

$$\mathbf{P}^{(ij)}\mathbf{P}^{(ij)} = \mathbf{I}$$

$$\mathbf{P}^{(12)}\mathbf{P}^{(12)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

A permutation matrix that pre-multiplies a matrix will permute the rows as we have seen. A permutation matrix that is post-multiplied will permute the columns.

$$\mathbf{A}\mathbf{P}^{(12)} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

$$\mathbf{P}^{(12)}\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

For part 1, using the definitions, we wish to show that:

$$\mathbf{L_k P^{(ij)}} = \mathbf{P^{(ij)}}(\mathbf{I} + \mathbf{P^{(ij)} m_k e_k^T})$$

By manipulating first the left hand side:

$$\begin{aligned}\mathbf{L_k P^{(ij)}} &= (\mathbf{I} + \mathbf{m_k e_k^T})\mathbf{P^{(ij)}} \\ &= \mathbf{P^{(ij)}} + \mathbf{m_k e_k^T P^{(ij)}})\end{aligned}$$

And then the right:

$$\begin{aligned}\mathbf{P^{(ij)}}(\mathbf{I} + \mathbf{P^{(ij)} m_k e_k^T}) &= \mathbf{P^{(ij)}} + \mathbf{P^{(ij)} P^{(ij)} m_k e_k^T}) \\ &= \mathbf{P^{(ij)}} + \mathbf{m_k e_k^T}\end{aligned}$$

We can see that it remains to show only that:

$$\mathbf{m_k e_k^T P^{(ij)}} = \mathbf{m_k e_k^T}$$

For part 2, use part 1.

3 HW 1.3 Hints

- The constants have to be independent of the particular vector x, though they can depend on the dimension.
- You should be able to solve this applying the definition of equivalence to the matrix norm equation and get the bounds you want. Think about how to make the matrix norm induced by vector norm B bigger than the matrix norm induced by vector norm A using the definition of vector norm equivalence.

4 The Method of Normal Equations

In class, we heard that the method of normal equations can be used to solve the linear least squares problem. To understand why this is the case, recall that we wish to minimize $||\mathbf{r}||_2^2 = ||\mathbf{b} - \mathbf{Ax}||_2^2$. So, let us define:

$$\begin{aligned}\phi(\mathbf{x}) &= ||\mathbf{r}||_2^2 \\ &= \mathbf{r}^T\mathbf{r} \\ &= (\mathbf{b} - \mathbf{Ax})^T(\mathbf{b} - \mathbf{Ax}) \\ &= \mathbf{b}^T\mathbf{b} - \mathbf{x}^T\mathbf{A}^T\mathbf{b} - \mathbf{b}^T\mathbf{Ax} + \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} \\ &= \mathbf{b}^T\mathbf{b} - 2\mathbf{x}^T\mathbf{A}^T\mathbf{b} + \mathbf{x}^T\mathbf{A}^T\mathbf{Ax}\end{aligned}$$

To find the minimia of this function, we first take its gradient

$$\nabla\phi(\mathbf{x}) = -2\mathbf{A}^T\mathbf{b} + 2\mathbf{A}^T\mathbf{Ax}$$

and then set it to 0, which gives us

$$\nabla\phi(\mathbf{x}) = \mathbf{0} = -2\mathbf{A}^T\mathbf{b} + \mathbf{A}^T\mathbf{Ax} \Rightarrow \mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$$

which is precisely the method of normal equations.