Math 25a Solution Set #8 (Part A)

Isidora Milin

December 2001

Problem 1

For an $n \times n$ matrix A, you were asked to show that TFAE ("the following are equivalent"):

(1) A has an inverse.

(2) The columns of A are LI.

(3) The rows of A are LI.

When proving TFAE-type statements it is customary (and a sign of good style!) to first lay out the plan of attack, such as (2) $\Rightarrow$ (1) $\Rightarrow$ (2) and (1) $\Rightarrow$ (3) $\Rightarrow$ (1) and then prove each implication in a separate paragraph, with the first sentence in each stating what you intend to prove in it. Math 25a isn't Expos (thank God!), but you should get used to making life easier for those honoured to read your proofs...

(2) $\Rightarrow$ (1). If columns of A are LI, since there are n of them, they span $\mathbb{R}^n$. Consider the linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$ whose matrix with respect to the standard basis on $\mathbb{R}^n$ is A. The columns of A are the images of the basis vectors under L. Since these span $\mathbb{R}^n$, L is surjective. But since L is a map between two spaces of the same dimension, we know this implies L is bijective and thus invertible. Invertibility of L implies that of A, which proves the implication.

(1) $\Rightarrow$ (2). If A has an inverse, A^{-1}, then if L, L' are lin. transf. with matrices (wrt. standard bases on $\mathbb{R}^n$) A and A^{-1} respectively, then LL' and $L'L$ are linear transformations whose matrix is the identity, so that $L' = L^{-1}$ and L is invertible. This implies that $im(L) = \mathbb{R}^n$. But we also know that $im(L)$ is the span of the columns of A, so that columns of A must span $\mathbb{R}^n$. Since there are n of them, they can span $\mathbb{R}^n$ only if they are LI, which proves this implication.

(1) $\Rightarrow$ (3). Recall from previous homework that if A is invertible, so is A^T, and $(A^T)^{-1} = (A^{-1})^T$. But know we can apply the implication above ((1) $\Rightarrow$ (2)), to conclude that columns of A^T are LI. But columns of A^T are just the rows of A, so rows of A are LI as well.

(3) $\Rightarrow$ (1). Similar "transpose trick" does the job here too – Since the rows of A are LI, it follows A^T has LI columns, so that by the first implication we have shown it has to be invertible. But then $A = (A^T)^T$ is invertible as well, with $A^{-1} = ((A^T)^{-1})^T$, and we are done.

The problem can be done using the fact that the matrix is invertible iff it row-reduces to the identity matrix (as many of you stated in your solutions), but as you see, a proof that doesn't mention row-reduction is possible too! ∎

Problem 2

Part (a). This was 2.7.3 in the textbook. The recurrent formula for the computation of the k-th root function via Newton's method is obtained from the general formula by elementary single variable calculus. The function whose zero we are "zeroing on" is $f(x) = x^k - b$, and let a_0 be our initial guess and g the iterative function we are looking for (so that $g(a_n) = a_{n+1}$). Then we have:

$$g(a_n) = a_n - \frac{f(a_n)}{f'(a_n)}$$

Now, in our case $f'(x) = kx^{k-1}$, so that we get:

$$a_{n+1} = g(a_n) = a_n - \frac{a_n^k - b}{ka_n^{k-1}} = \frac{(k-1)a_n^k + b}{ka_n^{k-1}} = \frac{1}{k}((k-1)a_n + \frac{b}{a_n^{k-1}})$$

The formula above can be interpreted as the average of k numbers, one of which is $\frac{b}{a_n^{k-1}}$, whereas the $k-1$ others all equal a_n. This is just the weighted average of a_n and $\frac{b}{a_n^{k-1}}$, with weights $k-1$ and 1, resp. In general, a weighted average of numbers $a_1, \ldots, a_n$, with weights $w_1, \ldots, w_n$, is the expression:

$$\frac{a_1 w_1 + \ldots + a_n w_n}{w_1 + \ldots + w_n}$$

∎

Part (b). This was 2.7.4(a) in the textbook. To apply Newton's method to the calculation of $9^{1/3}$ with six decimals and initial guess $a_0 = 2$, just plug in $k = 3$, $b = 9$ in the formula of the previous problem and start iterating by calculating $g(2)$, $g(g(2))$ where

$$g(a) = a - \frac{a^3 - 9}{3a^2}$$

. Most people obtained $a_3 \approx 2.080084$.

Part (c). This was 2.7.14$(a), (b)$ in the textbook.

The first part was to show that with $a_0 > 0$ and $b > 0$ in problem 2.7.3, the convergence of the sequence a_n is always to the (unique) positive root $\alpha = b^{1/k}$ of $f(x) = x^k - b$.

Since our f satisfies $f'(x) > 0$ and $f''(x) > 0$ (on $\mathbb{R}^+$), we have that f is increasing and concave on $\mathbb{R}^+$, so that for $x < \alpha \Rightarrow f(x) > 0$ and $x > \alpha \Rightarrow$

$f(x) > 0$. The concavity of f implies that if we start with $0 < a_0 < \alpha$ we get $a_1 > \alpha$, so that *wlog* we can assume $a_0 > \alpha$. From the iterative formula

$$a_{n+1} = g(a_n) = a_n - \frac{f(a_n)}{f'(a_n)}$$

and the above stated properties of f it follows that $\forall n \; a - n > 0 \Rightarrow a_{n+1} < a_n$. The weighted average formula from problem 2.7.3 does give us that all iterations of a positive number are positive, so that the sequence a_n is decreasing and bounded below by 0. It hence converges to its infimum, call it β. We know $\beta \geq 0$ since β is the greatest lower bound and 0 is a lower bound for a_n.

Noticing that

$$\beta = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} g(a_n) = g(\lim_{n\to\infty} a_n) = g(\beta)$$

gives us a condition on β:

$$\beta = \frac{1}{k}((k-1)\beta + \frac{b}{\beta^{k-1}}$$

from which we conclude that $\beta^k = b$. But $\beta \geq 0$ and $\beta^k = b$ implies $\beta = \alpha$, which is all we needed to show!

The second part was to show why this fails with the divide and average algorithm:

$$a_{n+1} = \frac{1}{2}(a_n + \frac{b}{a_n^{k-1}})$$

Define $h(x) = \frac{1}{2}(x + \frac{b}{x^{k-1}})$ so that the above algorithm gives a sequencea_n such that $a_{n+1} = h(a_n)$. Notice that if $0 < \alpha = b^{1/k}$, we have $h(\alpha) = \alpha$, and also that $h(x) = x \Rightarrow x^k = b$. Also, $h'(x) = \frac{1}{2}(1 - b(k-1)\frac{1}{x^k})$, so that $h'(\alpha) = \frac{2-k}{2}$. This means that for $k > 4$ we'll have $|h'(\alpha)| > 1$.

Now we are equipped to show that the divide and average algorithm fails for $k > 4$. Suppose, for the sake of contradiction, that $\lim_{n\to\infty} a_n$ exists. It has to be a fixed point of h, and for $a_0 > 0$ we get a sequence of positive numbers, so that the limit is ≥ 0 and hence $\lim_{n\to\infty} a_n = \alpha$. We know h' is a continuous function. Fix some $0 < \epsilon < |h'(\alpha)| - 1$ and choose δ such that $|x - \alpha| < \delta \Rightarrow |h'(x) - h'(\alpha)| < \epsilon$. Then we get $|x - \alpha| < \delta \Rightarrow |h'(x)| > 1$. Choose N such that $n \geq N \Rightarrow |a_n - \alpha| < \delta$. For $\forall n > N$ we then have (using the mean value theorem from single variable calculus):

$$|\alpha - a_n| = |h(\alpha) - h(a_{n-1})| = |h'(c)||\alpha - a_{n-1}| > |\alpha - a_{n-1}|$$

where the last inequality follows from:

$$c \in [\alpha, a_n] \Rightarrow |\alpha - c| \leq |\alpha - a_n| < \delta \Rightarrow |h'(c)| > 1$$

. We conclude that $\forall n > N$, $|\alpha - a_n| > |\alpha - a_{n-1}|$, so that the sequence a_n is not Cauchy and hence diverges. ▮