Additional Notes and Solution Manual For: Matrix Computations: Third Edition by Gene H. Golub and Charles F. Van Loan

John L. Weatherwax*

February 9, 2007

Chapter 2 (Matrix Analysis):

Basic Ideas from Linear Algebra

P 2.1.1 (existence of a p rank factorization of A)

Assume A is mxn and of rank r. The using elementary elimination matrices we can reduce A to its row reduced echelon form R, given by $E_1 A = R$ or $A = E_1^{-1} R = E_2 R$. In this reduction E_2 is mxm and R is mxn. Because R has $n - r$ zero rows we can block decompose it as follows

$$R = \begin{bmatrix} \hat{R}_{r \times n} \\ 0_{m-r \times n} \end{bmatrix}$$

where we have listed the dimensions of the the block matrices next to them. Now $\hat{R}$ is of rank r. In addition, block decomposing E_2 as

$$E_2 = \begin{bmatrix} \hat{E}_{m \times r} & \tilde{E}_{m \times m-r} \end{bmatrix} .$$

Now since E_2 is of rank m the first r columns of E_2 is a matrix of rank r. This block decomposition gives for A the expression

$$A = \begin{bmatrix} \hat{E} & \tilde{E} \end{bmatrix} \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} = \hat{E}\hat{R}$$

This gives the decomposition of A into $\hat{E}$ of size mxr and $\hat{R}$ of size rxn each of rank r, thus proving the decomposition.

*wax@alum.mit.edu

P 2.1.2 (the matrix product rule)

This result is basically a consequence of the definition of matrix multiplication. For example the ij-th element of the produce $A(\alpha)B(\alpha)$ is given by

$$\sum_{k=1}^{r} a_{ik}(\alpha)b_{kj}(\alpha)$$

where $a_{ik}(\alpha)$ is the ik-th element of A and $b_{kj}(\alpha)$ is the kj-th element of B. From this we then have that

$$\frac{d}{d\alpha}\sum_{k=1}^{r} a_{ik}(\alpha)b_{kj}(\alpha) = \sum_{k=1}^{r} \frac{da_{ik}(\alpha)}{d\alpha}b_{kj}(\alpha) + \sum_{k=1}^{r} a_{ik}(\alpha)\frac{db_{kj}(\alpha)}{d\alpha}$$

Which we recognize as the ij-th element of $\frac{dA}{d\alpha}B$ plus the ij-th element of $A\frac{dB}{d\alpha}$ proving the desired theorem.

P 2.1.3 (matrix inverse differentiation)

To show this consider the derivative of the expression

$$A(\alpha)A(\alpha)^{-1} = I$$

with respect to α.

$$\frac{d}{d\alpha}(A(\alpha)A(\alpha)^{-1}) = 0$$

Using the result of P 2.1.3 we have that

$$\frac{dA(\alpha)}{d\alpha}A(\alpha)^{-1} + A(\alpha)\frac{dA(\alpha)^{-1}}{d\alpha} = 0$$

and solving for $\frac{dA(\alpha)^{-1}}{d\alpha}$ we have that

$$\frac{dA(\alpha)^{-1}}{d\alpha} = -A(\alpha)^{-1}\frac{dA(\alpha)^{-1}}{d\alpha}A(\alpha)^{-1}$$

the desired result.

P 2.1.4 (the gradient of the matrix inner product)

We desire the gradient of the function

$$\phi(x) = \frac{1}{2}x^T Ax - x^T b\,.$$

The i-th component of the gradient is given by

$$\frac{\partial \phi}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{1}{2}x^T A x - x^T b\right) = \frac{1}{2}e_i^T A x + \frac{1}{2}x^T A e_i - e_i^T b$$

where e_i is the i-th elementary basis function for $\mathbb{R}^n$, i.e. it has a 1 in the i-th position and zeros everywhere else. Now since

$$(e_i^T A x)^T = x^T A^T e_i^T = e_i^T A x\,,$$

the above becomes

$$\frac{\partial \phi}{\partial x_i} = \frac{1}{2}e_i^T A x + \frac{1}{2}e_i^T A^T x - e_i^T b = e_i^T\left(\frac{1}{2}(A + A^T)x - b\right)\,.$$

Since multiplying by e_i^T on the left selects the i-th row from the expression to its right we see that the full gradient expression is given by

$$\nabla \phi = \frac{1}{2}(A + A^T)x - b\,,$$

as requested in the text. Note that this expression can also be proved easily by writing each term in components.

P 2.1.5 (solutions to rank one updates of A)

If x solves $(A + uv^T)x = b$, by the Sherman-Morrison-Woodberry formula (equation 2.1.4 in the book), with U and V vectors with n components we have x given by

$$\begin{aligned} x &= (A^{-1} - A^{-1}u(I + v^T A^{-1}u)^{-1}v^T A^{-1})b \\ &= A^{-1}b - A^{-1}u(I + v^T A^{-1}u)^{-1}v^T A^{-1}b\,. \end{aligned}$$

Since u and v are vectors the expression $v^T A^{-1}u$ is a *scalar* and the I is also a scalar namely the number 1. Multiplying the above by A on the left the linear system that x must satify

$$Ax = b - u(1 + v^T A^{-1}u)^{-1}v^T A^{-1}b\,.$$

In this expression, both $v^T A^{-1}u$ and $v^T A^{-1}b$ are scalars, thus by factoring out the only vector u the above is equivalent to

$$Ax = b - \left(\frac{v^T A^{-1}b}{(1 + v^T A^{-1}u)}\right)u\,.$$

Therfore x is the solution to a modified system given by $Ax = b + \alpha u$ with α given by

$$\alpha = -\left(\frac{v^T A^{-1}b}{1 + v^T A^{-1}u}\right)\,.$$