Study Questions Homework 3 Foundations of Computational Math 1 Fall 2021

Problem 3.1

Let x and y be two vectors in $\mathbb{R}^n$.

3.1.a. Show that given x and y the value of $\|x - \alpha y\|_2$ is minimized when

$$\alpha_{min} = \frac{x^T y}{y^T y}$$

3.1.b. Show that $x = y\alpha_{min} + z$ where $y^T z = 0$, i.e., x is easily written as the sum of two orthogonal vectors with specifed minimization properties.

Problem 3.2

Recall that an elementary reflector has the form $Q = I + \alpha z z^T \in \mathbb{R}^{n \times n}$ with $\|z\|_2 \neq 0$.

3.2.a. Show that Q is orthogonal if and only if

$$\alpha = \frac{-2}{z^T z} \quad \text{or} \quad \alpha = 0$$

3.2.b. Given $v \in \mathbb{R}^n$, let $\gamma = \pm\|v\|$ and $z = v + \gamma e_1$. Assuming that $z \neq v$ show that

$$\frac{z^T z}{z^T v} = 2$$

3.2.c. Using the definitions and results above show that $Qv = -\gamma e_1$

Problem 3.3

Let $x \in \mathbb{R}^n$ be a known vector with components $\xi_i = e_i^T x,\ 1 \leq i \leq n$ and consider the computation of

$$\nu = \xi_1 - \|x\|_2$$

where $\|x\|_2^2 = \sum_{i=1}^n \xi_i^2$. (Recall this is a key computation in the production of a Householder reflector in least squares problems.) When $\xi_1 > 0$ and $\xi_1 \approx \|x\|_2$ the cancellation in the subtraction may result in a significant loss of accuracy.

Find an alternate expression for ν that does not suffer from cancellation when $\xi_1 > 0$ and $\xi_1 \approx \|x\|_2$. (**Hint:** Consider a difference of squares.)

Problem 3.4

Consider a Householder reflector, H, in $\mathbb{R}^2$. Show that

$$H = \begin{pmatrix} -\cos(\phi) & -\sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$$

where ϕ is some angle.

Problem 3.5

Suppose you are given the nonsingular tridiagonal matrix $T \in \mathbb{R}^{n\times n}$ For example, if $n = 6$ then

$$\begin{pmatrix} \alpha_1 & \beta_1 & 0 & 0 & 0 & 0 \\ \gamma_2 & \alpha_2 & \beta_2 & 0 & 0 & 0 \\ 0 & \gamma_3 & \alpha_3 & \beta_3 & 0 & 0 \\ 0 & 0 & \gamma_4 & \alpha_4 & \beta_4 & 0 \\ 0 & 0 & 0 & \gamma_5 & \alpha_5 & \beta_5 \\ 0 & 0 & 0 & 0 & \gamma_6 & \alpha_6 \end{pmatrix}$$

3.5.a Suppose you use Householder reflectors to transform T to upper triangular, i.e.,

$$H_{n-1} \dots H_1 T = R.$$

What is the zero/nonzero structure of R?

3.5.b What is the structure of each of the reflectors H_i?

3.5.c Let $T^{(i)} = H_i H_{i-1} \dots H_1 T$. What is the structure of $T^{(i)}$?

3.5.d What is the computational complexity of the factorization, i.e., what is k in $O(n^k)$? (You do not have to determine the constant in the complexity expression.)

To answer the questions on structure, in addition to characterizing it algebraically, include a $*$, 0 matrix pattern for $n = 6$ to make it clear.

Problem 3.6

3.6.a

This part of the problem concerns the computational complexity question of operation count.

For both LU factorization and Householder reflector-based orthogonal factorization, we have used elementary transformations, T_i, that can be characterized as rank-1 updates to the identity matrix, i.e.,

$$T_i = I + x_i y_i^T, \quad x_i \in \mathbb{R}^n \text{ and } y_i \in \mathbb{R}^n$$

Gauss transforms and Householder reflectors differ in the definitions of the vectors x_i and y_i. Maintaining computational efficiency in terms of a reasonable operation count usually implies careful application of associativity and distribution when combining matrices and vectors.

Suppose we are to evaluate

$$z = T_3T_2T_1v = (I + x_3y_3^T)(I + x_2y_2^T)(I + x_1y_1^T)v$$

where $v \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$. Show that by using the properties of matrix-matrix multiplication and matrix-vector multiplication, the vector z can be evaluated in $O(n)$ computations (a good choice of version for an algorithm) or $O(n^2)$ computations (a bad choice of version for an algorithm) or $O(n^3)$ computations (a very bad choice of version for an algorithm).

3.6.b

This part of the problem concerns the computational complexity question of storage space.

Recall, that we discussed and programmed an **in-place** implementation of LU factorization that was very efficient in terms of storage space. An array with n^2 entries initialized with $array(I, J) = \alpha_{ij}$ could be used to store the n^2 entries needed to specify L and U, i.e., λ_{ij} for $j < i$, $2 \leq i \leq n$ and $1 \leq j \leq n - 1$, and μ_{ij} for $i < j$, $1 \leq i \leq n$ and $1 \leq j \leq n$.

Let $A \in \mathbb{R}^{n \times k}$, $n \geq k$, and $rank(A) = k$. Consider the use of Householder reflectors, H_i, $1 \leq i \leq k$, to transform A to upper trapezoidal form, i.e.,

$$H_kH_{k-1} \cdots H_2H_1A = \begin{pmatrix} R \\ 0 \end{pmatrix}$$

$$R \in \mathbb{R}^{k \times k} \text{ nonsingular upper triangular}$$

Suppose you are given an array with $n \times k$ entries initialized with $array(I, J) = \alpha_{ij}$ and you are to implement your algorithm using minimal storage.

(i) Are you able to store all of the information needed to specify the H_i, $1 \leq i \leq k$ and R within the array with $n \times k$ entries? Justify your answer.

(ii) If you are not able to store all of the information in the array, how much extra storage do you need and what do you store in it?

Problem 3.7

3.7.a

Consider a set $\{x_i\}$, $i = 1, \ldots n$, of vectors $x_i \in \mathbb{R}^d$ and the vector 2-norm defined by the standard Euclidean inner product, i.e., $\langle x, y \rangle = x^Ty$ and $\|x\|_2^2 = \langle x, x \rangle$. The Karcher mean or intrinsic mean of the vectors with respect to this norm is the unique minimizer of

$$f(x) = \min_{x \in \mathbb{R}^d} \sum_{i=1}^{n} \|x - x_i\|_2^2$$

and has a well-known closed form given by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Now consider the inner product $\langle x, y \rangle_M = x^T M y$ and the norm it defines $\|x\|_M^2 = \langle x, x \rangle_M$, where M is a symmetric positive definite matrix. The Karcher mean or intrinsic mean, $\tilde{x}$, of the vectors with respect to this M-norm is the unique minimizer of

$$f_M(x) = \min_{x \in \mathbb{R}^d} \sum_{i=1}^{n} \|x - x_i\|_M^2$$

Determine a closed form expression for $\tilde{x}$ and the effect of M, if any, on the mean. Justify your answer.

3.7.b

Suppose $b \in \mathbb{R}^d$ is a known vector. Consider the unit sphere defined by the M-norm

$$\mathcal{S} = \{x \in \mathbb{R}^d \mid \|x\|_M^2 = 1\}.$$

Determine the approximating vector $\hat{x}$ that solves

$$g(x) = \min_{x \in \mathcal{S}} \|x - b\|_M^2.$$

Justify your answer.