Problem

Let Q denote an $n \times n$ real orthogonal matrix. Then $Q^T Q = QQ^T = I$.

(1) Show that the linear transformation by Q does not change the length of a vector.

(2) Show that the determinant of Q is either -1 or 1.

(3) In the n dimensional real space, consider a line whose direction vector is unit vector v. Point y is the mirror image of point x with respect to the line. Namely, the middle point of the line segment connecting x and y is included in the line and the line segment is orthogonal to the line.

(3.1) Describe the linear transformation from x to y as $y = Qx$.

(3.2) Show that Q is an orthogonal matrix.

Nucleotide sequences of protein coding regions of mRNAs are translated into amino acids corresponding to the codons of three nucleotides. An amino acid often corresponds to multiple kinds of codons of three nucleotides; there are six, four, two, and three kinds of codons for Arg (Arginine), Ala (Alanine), Asp (Aspartic acid), and Ile (Isoleucine) respectively.

There are given 130 sequences, which are 12 consecutive nucleotides in protein coding regions of mRNAs corresponding to four consecutive amino acids, Arg, Ala, Asp and Ile in the translated proteins. Assume that those 130 sequences are mutually different.

The result of counting the occurrence of codons for each amino acid in the 130 sequences is shown in the following table.

amino acid	Arg						Ala				Asp		Ile		
codon	AGA	AGG	CGA	CGC	CGG	CGU	GCA	GCC	GCG	GCU	GAC	GAU	AUA	AUC	AUU
count	21	22	23	24	18	22	33	30	35	32	68	62	48	40	42

(1) How many kinds of possible nucleotide sequences of 12 nucleotides correspond to those four amino acids in that order?

(2) Explain briefly an efficient algorithm that sorts those 130 sequences in alphabetical order (A to Z).

(3) Assume that there are the following two sequences in those 130 sequences.

CGCGCAGACAUA

CGAGCAGACAUC

Answer the sorted rank of those two sequences in the 130 sequences.