William Mardick-Kanter

1.6 *Vector of differences.* Suppose x is an n-vector. The associated vector of differences is the $(n-1)$-vector d given by $d = (x_2 - x_1, x_3 - x_2, \ldots, x_n - x_{n-1})$. Express d in terms of x using vector operations (*e.g.*, slicing notation, sum, difference, linear combinations, inner product). The difference vector has a simple interpretation when x represents a time series. For example, if x gives the daily value of some quantity, d gives the day-to-day changes in the quantity.

- Slicing Notation

  let i and k be integers such that $0 < i < k < n$

  $d_{i:k} = (x_{i+1} - x_i, \ldots, x_k - x_{k-1})$

- Difference

  $x_{2:4} = (x_2, x_3, x_4)$ $\quad d = x_{2:4} - x_{1:3}$

  $x_{1:3} = (x_1, x_2, x_3)$

  $d = x_{2:n} - x_{1:n-1}$

1.7 *Transforming between two encodings for Boolean vectors.* A Boolean n-vector is one for which all entries are either 0 or 1. Such vectors are used to encode whether each of n conditions holds, with $a_i = 1$ meaning that condition i holds. Another common encoding of the same information uses the two values -1 and $+1$ for the entries. For example the Boolean vector $(0, 1, 1, 0)$ would be written using this alternative encoding as $(-1, +1, +1, -1)$. Suppose that x is a Boolean vector with entries that are 0 or 1, and y is a vector encoding the same information using the values -1 and $+1$. Express y in terms of x using vector notation. Also, express x in terms of y using vector notation.

$y_1 = -1 \qquad x_1 = 0$

$2(0) - 1 = -1$

$2(1) - 1 = 2 - 1 = 1$

$y_i = 2x_i - 1 \qquad\qquad x_i = \left(\frac{y_i + 1}{2}\right)$

1.13 *Average age in a population.* Suppose the 100-vector x represents the distribution of ages in some population of people, with x_i being the number of $i-1$ year olds, for $i = 1, \ldots, 100$. (You can assume that $x \neq 0$, and that there is no one in the population over age 99.) Find expressions, using vector notation, for the following quantities.

(a) The total number of people in the population.

(b) The total number of people in the population age 65 and over.

(c) The average age of the population. (You can use ordinary division of numbers in your expression.)

Total number of people:

- To find the total population we would sum the elements of the vector

$$\sum_{i=1}^{100} x_i = (x_1 + x_2 + \cdots x_{100}) \quad || \quad X^T 1_{100}$$

This notation gives every element multiplied by 1, then summed $X^T 1_{100} = X_1(1) + X_2(1) + \cdots X_{100}(1)$

Total number of people 65 and older

- We sum starting at $i=66$ and continue to $i=100$

$$\sum_{i=66}^{100} x_i = (X_{66} + X_{67} + \cdots X_{100}) \quad || \quad X_{66:100}^T 1_{35}$$

Vector notation for equivalent summation

To find the Average age we will use a weighted average

We will multiply each index (which contains the number of people of i-1 age)

Let n be the size of the age vector

Let a be an n-vector with elements such that $a_i = i-1$

If we transform $a^T x$ we will get the summed ages of every person in our population so to get the average age we need to divide this by population size which is given by transforming a 1_{100} vector with X

let $p = X^T 1_{100}$

$$\frac{1}{p} \overset{a}{\begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \\ 99 \end{bmatrix}}^T \overset{X}{\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ \vdots \\ X_{100} \end{bmatrix}} = \frac{1}{X^T 1_{100}} (a^T x)$$

1.16 *Inner product of nonnegative vectors.* A vector is called *nonnegative* if all its entries are nonnegative.

(a) Explain why the inner product of two nonnegative vectors is nonnegative.

(b) Suppose the inner product of two nonnegative vectors is zero. What can you say about them? Your answer should be in terms of their respective sparsity patterns, *i.e.*, which entries are zero and nonzero.

a) If a is a nonnegative vector and b is a nonnegative vector
$\forall i$ s.t. $0 < i \leq n$ $a_i > 0$, $b_i > 0$
since the product of two non negative numbers is always non negative
then $a^T b$ can't be negative

Example $a_5 = \{1, 2, 3, 4, 5\}$
$b_5 = \{0, 0, 1, 0, 0\}$

$$a^T b = (0 \cdot 1) + (0 \cdot 2) + (1 \cdot 3) + (0 \cdot 4) + (0 \cdot 5)$$
$$= 3$$

b) using the previous example, if the two vectors don't have indices s.t $a_i > 0$ $b_i > 0$ then the inner product could be 0 at the least

1.18 *Linear combinations of linear combinations.* Suppose that each of the vectors $b_1, \ldots, b_k$ is a linear combination of the vectors $a_1, \ldots, a_m$, and c is a linear combination of $b_1, \ldots, b_k$. Then c is a linear combination of $a_1, \ldots, a_m$. Show this for the case with $m = k = 2$. (Showing it in general is not much more difficult, but the notation gets more complicated.)

Max size 2 for simplicity, let b
$b_1 = \{\beta_1 a_1, \beta_2 a_2\}$
$b_2 = \{\gamma_1 a_1, \gamma_2 a_2\}$
And c,
$$c = (\varepsilon_1 b_1, \varepsilon_2 b_2)$$
$$= \varepsilon_1(\beta_1 a_1 + \beta_2 a_2) + \varepsilon_2(\gamma_1 a_1 + \gamma_2 a_2)$$

Using distributive property we can distribute scalars

$$= \varepsilon_1 \beta_1 a_1 + \varepsilon_1 \beta_2 a_2 + \varepsilon_2 \gamma_1 a_1 + \varepsilon_2 \gamma_2 a_2$$
$$= a_1(\varepsilon_1 \beta_1 + \varepsilon_2 \gamma_1) + a_2(\varepsilon_1 \beta_2 + \varepsilon_2 \gamma_2)$$

since these are just scalars, they will produce scalars $(\varepsilon_1 \beta_2 + \varepsilon_2 \gamma_2)$ $(\varepsilon_1 \beta_1 + \varepsilon_2 \gamma_1)$

$$= \phi_1 a_1 + \phi_2 a_2$$

let $\phi_1 = (\varepsilon_1 \beta_1 + \varepsilon_2 \gamma_1)$
$\phi_2 = (\varepsilon_1 \beta_2 + \varepsilon_2 \gamma_2)$

which is a linear combination of vector a.