Math review exercises

Jitkomut Songsiri

Department of Electrical Engineering
Chulalongkorn University

July 23, 2023

Notes to readers. These exercises are in undergrad-level (year 1-2) and meant to be a review before students taking any upper-undergrad or graduate classes in control systems, and optimization. You are expected to solve these exercises by yourself (and by hand).

Contents

1 Notation

notation	**description**
$\mathbf{R}$	set of real numbers
$\mathbf{R}^n$	set of real vectors of length n
$\mathbf{R}^{m\times n}$	set of real matrices of size $m\times n$
$\mathbf{C}$	set of complex numbers
$\mathbf{C}^n$	set of complex vectors of length n
$\mathbf{C}^{m\times n}$	set of complex matrices of size $m\times n$
$\mathbf{S}^n$	set of symmetric matrices of size $n\times n$
$\mathcal{N}(T)$	nullspace of linear transformation T
$\mathcal{R}(T)$	range space of linear transformation T
$\mathbf{cov}(X)$	covariance matrix of X
$\nabla f, \nabla^2 f$	gradient and Hessian of $f : \mathbf{R}^n \to \mathbf{R}$

A column vector is denoted by $x \in \mathbf{R}^n$ where x_i is the ith element of x. A regtangular matrix $A \in \mathbf{R}^{m\times n}$ is denoted by a capital letter where a_{ij} is the (i,j) entry of A. The matrix A^T is the transpose of A. We can partition A and B into column blocks and row blocks, respectively.

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1^T \\ \vdots \\ b_m^T \end{bmatrix}.$$

We denote $\mathbf{e}_k$ a standard unit vector in $\mathbf{R}^n$, *e.g.*, $\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, 0, \ldots, 0)$. An all-one vector is denoted as $\mathbf{1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$. We can construct a diagonal matrix from a vector using a notation of $\mathbf{diag}(x) = \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}$.

The 2-norm of a vector x is denoted by $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ and it should be clear that $\|x\|_2^2 = x^T x$.

2 Matrix and vector

2.1 Multiplications

Perform the following multiplications in 1 minute each.

$$\begin{bmatrix} -2 & 3 & 3 \\ 0 & -2 & 0 \\ 3 & 3 & 2 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 4 \end{bmatrix} = \tag{1}$$

$$\begin{bmatrix} -3 & -1 & 4 \end{bmatrix} \begin{bmatrix} -2 & 3 & 3 \\ 0 & -2 & 0 \\ 3 & 3 & 2 \end{bmatrix} = \tag{2}$$

$$\begin{bmatrix} -3 & -1 & 4 \end{bmatrix} \begin{bmatrix} -2 & 3 & 3 \\ 0 & -2 & 0 \\ 3 & 3 & 2 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 4 \end{bmatrix} = \tag{3}$$

$$\begin{bmatrix} -2 & 3 & 3 \\ 0 & -2 & 0 \\ 3 & 3 & 2 \end{bmatrix} \begin{bmatrix} -3 & -1 \\ -1 & 1 \\ 4 & -3 \end{bmatrix} = \tag{4}$$

$$\begin{bmatrix} -3 & -1 & 4 \\ -1 & 1 & -3 \end{bmatrix} \begin{bmatrix} -2 & 3 & 3 \\ 0 & -2 & 0 \\ 3 & 3 & 2 \end{bmatrix} = \tag{5}$$

$$\begin{bmatrix} 1 & 1 & -3 \\ 2 & 1 & 0 \\ -2 & -2 & 3 \end{bmatrix} \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{bmatrix} = \tag{6}$$

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & -3 \\ 2 & 1 & 0 \\ -2 & -2 & 3 \end{bmatrix} = \tag{7}$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E \\ F \end{bmatrix} = \tag{8}$$

$$(A_{ij}\text{'s are matrices}) \quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{33} \end{bmatrix}^T \begin{bmatrix} B & C \\ D & E \\ F & G \end{bmatrix} = \tag{9}$$

2.2 Get it right

Write these expressions in terms of matrix/vector components explicitly. For example, when A is partitioned in column blocks,

$$A\,\mathbf{diag}(x) = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix} = \begin{bmatrix} x_1 a_1 & x_2 a_2 & \cdots & x_n a_n \end{bmatrix}$$

In this section, we assume A partitioned in columns, while B is partitioned in rows. The vectors x and y have dimensions that agree with the context.

$$
\begin{aligned}
xy^T &= && (10)\\
xx^Tyy^T &= && (11)\\
A^T &= && (12)\\
A^TA &= && (13)\\
AA^T &= && (14)\\
B^T &= && (15)\\
B^TB &= && (16)\\
BB^T &= && (17)\\
\mathbf{e}_1\mathbf{e}_1^T &= && (18)\\
\mathbf{e}_2\mathbf{e}_5^T &= && (19)\\
\mathbf{e}_j^T\mathbf{e}_j &= && (20)\\
C\mathbf{e}_k &= && (21)\\
\mathbf{e}_1^TC\mathbf{e}_1 &= && (22)\\
\mathbf{e}_i^TC\mathbf{e}_j &= && (23)\\
\mathbf{1}^Tx &= && (24)\\
A\mathbf{1} &= && (25)\\
\mathbf{1}^TB &= && (26)\\
\mathbf{1}^TC\mathbf{1} &= && (27)\\
A\,\mathbf{diag}(x) &= && (28)\\
\mathbf{diag}(x)B &= && (29)
\end{aligned}
$$

2.3 Common expressions

Expand the following expressions in terms of x_i's, y_i's and a_{ij}'s where $x, y \in \mathbf{R}^3$ and $A = [a_{ij}] \in \mathbf{R}^3$ (10 minutes).

$$
\begin{aligned}
(x+y)^T(x+y) &= && (30)\\
\|x-y\|_2^2 &= && (31)\\
x^TAx &= && (32)\\
x^T(A+A^T)x &= && (33)\\
\frac{x^T(A+A^T)x}{2} &= && (34)
\end{aligned}
$$

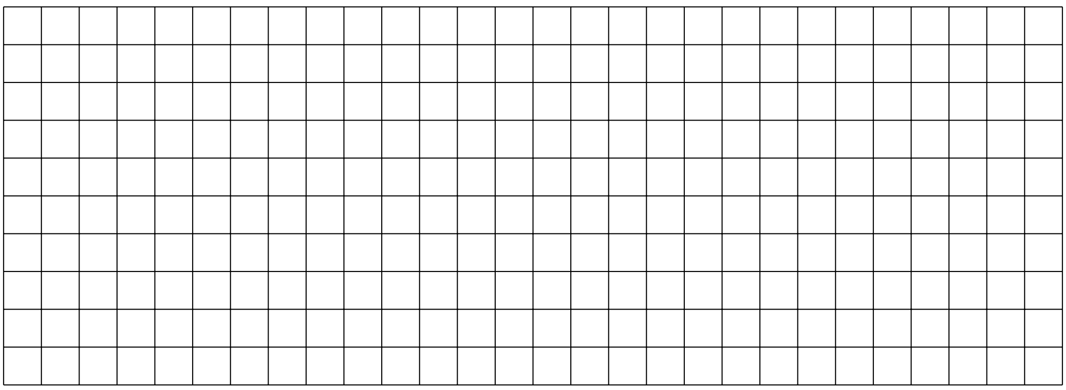
2.4 Represent vectors in $\mathbf{R}^2$ plane

Consider

$$x = (0,1), \quad y = (-4,2), \quad z = (2,3), \quad \alpha_1 = 0.5, \quad \alpha_2 = 1$$

Draw the resulting vectors in the diagram (10 minutes).

$$x, \quad -\alpha_1 y, \quad x - \alpha_1 y, \quad -\alpha_2 z, \quad (x-\alpha_1 y) - \alpha_2 z \tag{35}$$



2.5 Eigenvalues and eigenvectors

Compute eigenvalues and eigenvectors of the following matrices (by hand).

$$
\begin{aligned}
A_1 &= \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix} && (36)\\
A_2 &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} && (37)\\
A_3 &= \begin{bmatrix} 3 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} && (38)\\
A_4 &= \begin{bmatrix} A_1 & 0 \\ 0 & A_3 \end{bmatrix} && (39)\\
A_5 &= A_1^2 - 3A_1 + 2I && (40)\\
A_6 &= 3A_2^2 - 3A_3^2 + 4A_3 && (41)
\end{aligned}
$$

Review properties of eigenvalues, and relations of $\det(A), \mathbf{tr}(A)$ to eigenvalues of A.

2.6 Nullspace and range space

Find bases for the nullspace and range space, and their dimensions of the following matrices.

$$
A_1 = \begin{bmatrix} 0 & 0 & -2 & -1 & 1 \\ 1 & 0 & -1 & 0 & -1 \\ -1 & -2 & 1 & -2 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 3 & 1 & -1 \\ -5 & -2 & 1 \\ -3 & -2 & -1 \end{bmatrix}
$$

If we express a general solution to $Ax = 0$ (nullspace of A) as $x = Fz$ where $F \in \mathbf{R}^{n \times r}$ and $z \in \mathbf{R}^r$. Determine what F should be and the dimension r.

3 Calculus

3.1 Derivatives

Review the concept of first and second derivatives of function $f : \mathbf{R}^n \rightarrow \mathbf{R}$ (multivariate). Derive the gradient and Hessian of the following functions (30 minutes).

$$
\begin{aligned}
f(x) &= a(x_2 - x_1^2)^2 + b(1 - x_1)^2, \quad a, b \text{ are fixed} && (42)\\
f(x) &= \sum_{i=1}^{n} x_i \log(x_i) && (43)\\
f(x) &= a^T x, \quad a \in \mathbf{R}^n \text{ is fixed} && (44)\\
f(x) &= (1/2) x^T P x, \quad P \text{ is symmetric and fixed} && (45)\\
f(x) &= \frac{1}{1 + e^{-a^T x}} && (46)\\
f(x) &= e^{w^T x} - w^T x && (47)
\end{aligned}
$$

3.2 Regions in $\mathbf{R}^2$

Hand sketch the region of $x \in \mathbf{R}^2$ described by the following expressions. You may verify your results with using computer.

1. $2x_1 - 3x_2 = 6$
2. $2x_1 - 3x_2 \geq 6$
3. $x_1, x_2 \geq 0$ and $2x_1 + x_2 \leq 6$
4. $2x_1 + x_2 \leq 6$ and $2x_1 + x_2 \geq 2$

5. $2x_1^2 + x_2^2 \leq 2$
6. $2x_1^2 - 2x_1x_2 + x_2^2 \leq 2$
7. $2x_1^2 - 12x_1 - x_2 + 19 = 0$
8. $2x_1^2 - 12x_1 - x_2 + 19 \leq 0$
9. $|x_1| + |x_2| = c$, where $c > 0$. What happen when c is larger ?
10. $\max(0, 3x_1 - 2x_2) = c$, where $c \geq 0$
11. $\max(0, 3x_1 - 2x_2) \leq c$, where $c \geq 0$. What happen when c is larger ?
12. region that contains all the vectors orthogonal to the plane $\{x \mid x_1 - 4x_2 = c\}$ where $c = 0$ and $c = 5$.

3.3 Regions in $\mathbf{R}^3$

Use **computer** to plot the surface and contour of the following functions.

1. $f(x) = 2x_1 - 4x_2$, $x_0 = (1, 1)$
2. $f(x) = x_1^2 + x_2^2$, $x_0 = (2, 3)$
3. $f(x) = 2x_1^2 - 2x_1x_2 + x_2^2$, $x_0 = (-1, 1)$

For each function, compute the gradient of f and evaluate at x_0, and you will obtain $\nabla f(x_0)$ as a vector in $\mathbf{R}^2$. Create a hyperplane described by equation

$$0 = f(x_0) + \nabla f(x_0)^T(x - x_0) \quad \triangleq 0 = c + a_1x_1 + a_2x_2$$

where you notice that $(a_1, a_2) = \nabla f(x_0)$. When you include this hyperplane in the surface plot of f, this plane is supposed to touch the surface of f at x_0, because the hyperplane equation is just the first-order Taylor approximation of f at x_0. Notice the direction of $\nabla f(x_0)$ and its relation with the hyperplane.