SOLUTIONS MANUAL FOR

Numerical Methods and Optimization: Solutions and Exercises

by

Sergiy Butenko and Panos M. Pardalos

CRC Press
Taylor & Francis Group

SOLUTIONS MANUAL FOR

Numberical Methods and Optimization: Solutions to Exercises

by

Sergiy Butenko and Panos M. Pardalos

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an informa business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper
Version Date: 20141015

International Standard Book Number-13: 978-1-4665-7780-0 (Ancillary)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Preface

This solution manual is a work in progress. Solutions to the following exercises are missing as of now: 2.9, 3.1–3.14, 4.1–4.4, 4.6–4.14, 4.18–21, 5.1, 5.3, 5.4, 7.2–7.5, 8.2, 8.3, 8.4, 8.6, 8.8, 8.10, 8.12, 8.14–8.17, 9.12, 10.9, 11.7, 12.10, 12.11, 13.1, 13.3, 13.4, 13.9–13.14, 13.16, 14.2, 14.4, 14.5, 14.7, 14.9, 14.11, 14.13, 14.15. We expect to have these ready before January 1, 2015. Please report any typos/errors to Sergiy Butenko (butenko@tamu.edu). Thank you for your patience and support.

Sergiy Butenko and Panos Pardalos
October 12, 2014

Contents

Part I

Basics

Chapter 1

Preliminaries

1.1. Let $f : X \to Y$ be an arbitrary mapping and $X', X'' \subseteq X$, $Y', Y'' \subseteq Y$. Prove that

(a) $f^{-1}(Y' \cup Y'') = f^{-1}(Y') \cup f^{-1}(Y'')$;
(b) $f^{-1}(Y' \cap Y'') = f^{-1}(Y') \cap f^{-1}(Y'')$;
(c) $f(X' \cup X'') = f(X') \cup f(X'')$;
(d) $f(X' \cap X'')$ may not be equal to $f(X') \cap f(X'')$.

Solution:

(a) Consider $x \in f^{-1}(Y' \cup Y'')$. Then there exists $y \in Y' \cup Y''$ such that $f(x) = y$. This implies that $x \in f^{-1}(Y') \cup f^{-1}(Y'')$. Thus,

$$f^{-1}(Y' \cup Y'') \subseteq f^{-1}(Y') \cup f^{-1}(Y''). \tag{1.1}$$

Now, consider $x \in f^{-1}(Y') \cup f^{-1}(Y'')$. Then $x \in f^{-1}(Y')$ or $x \in f^{-1}(Y'')$. Since $f^{-1}(Y') \subseteq f^{-1}(Y' \cup Y'')$ and $f^{-1}(Y'') \subseteq f^{-1}(Y' \cup Y'')$, we have $x \in f^{-1}(Y' \cup Y'')$. Thus

$$f^{-1}(Y') \cup f^{-1}(Y'') \subseteq f^{-1}(Y' \cup Y''). \tag{1.2}$$

From (1.1) and (1.2) we have $f^{-1}(Y' \cup Y'') = f^{-1}(Y') \cup f^{-1}(Y'')$.

(b) Consider $x \in f^{-1}(Y' \cap Y'')$. Then there exists $y \in Y' \cap Y''$ such that $f(x) = y$. This implies that $x \in f^{-1}(Y') \cap f^{-1}(Y'')$. Thus,

$$f^{-1}(Y' \cap Y'') \subseteq f^{-1}(Y') \cap f^{-1}(Y''). \tag{1.3}$$

Now, consider $x \in f^{-1}(Y') \cap f^{-1}(Y'')$. Then $x \in f^{-1}(Y')$ and $x \in f^{-1}(Y'')$. Hence, there exist $y' \in Y'$ and $y'' \in Y''$ such that $f(x) = y'$ and $f(x) = y''$. Since $f(x)$ is unique, this implies that $y' = y'' \in Y' \cap Y''$. So, $x \in f^{-1}(Y' \cap Y'')$ and

$$f^{-1}(Y') \cap f^{-1}(Y'') \subseteq f^{-1}(Y' \cap Y''). \tag{1.4}$$

From (1.3) and (1.4) we have $f^{-1}(Y' \cap Y'') = f^{-1}(Y') \cap f^{-1}(Y'')$.

(c)

$$y \in f(X' \cup X'')$$
$$\Updownarrow$$
$$\text{there exists } x \in X' \text{ or } X'' \text{ such that } f(x) = y$$
$$\Updownarrow$$
$$y \in f(X') \cup f(X'').$$

(d) Consider, for example, $f(x) = x^2, X' = [-1, 0], X'' = [0, 1]$. Then $f(X') = f(X'') = f(X') \cap f(X'') = [0, 1]$, however, $f(X' \cap X'') = f(\{0\}) = \{0\}$.

1.2. Prove that the following sets are countable:

(a) the set of all odd integers;

(b) the set of all even integers;

(c) the set $2, 4, 8, 16, \ldots, 2^n, \ldots$ of powers of 2.

Solution: We have the following bijections with the countable set $\mathbb{Z}$ of all integers or $\mathbb{Z}_+$ of all positive integers:

(a) $f(n) = 2n + 1, n \in \mathbb{Z}$;

(b) $f(n) = 2n, n \in \mathbb{Z}$;

(c) $f(n) = 2^n, n \in \mathbb{Z}_+$.

1.3. Show that

(a) every infinite subset of a countable set is countable;

(b) the union of a countable family of countable sets $A_1, A_2, \ldots$ is countable;

(c) every infinite set has a countable subset.

Solution:

(a) Let A be the countable set, and let B be its infinite subset. Since A is countable, there is a bijection $f : \mathbb{Z}_+ \to A$. Let $f(n) = a_n \in A$ for any $n \in \mathbb{Z}_+$. We build a bijection $g : \mathbb{Z}_+ \to B$ such that $g(k) = b_k, k \geq 1$ as follows. Let b_k be the element of $\{a_n : n \geq 1\}$ with the k^{th} smallest index among the elements of A that belong to B. Since B is infinite, there is such an element for any $k \geq 1$. On the other hand, each element of B is one of the elements of $\{a_n : n \geq 1\}$ (since $B \subseteq A = \{a_n : n \geq 1\}$, so for any $b \in B$ there exists k such that $b_k = b$. Thus g is a bijection and B is countable.

(b) We can assume that no two sets have any elements in common (otherwise, we can consider $A_1, A_2 \backslash A_1, A_3 \backslash (A_1 \cup A_2), \ldots$, each of which is countable as a subset of a countable set).

We can write the elements of $A_1, A_2, \ldots$ in the form of an infinite table as follows:

$$\begin{array}{cccc} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

where a_{ij} is the j^{th} element of A_i, i, j=1,2, ... Clearly, this table contains all the elements of all the sets. We can count the elements of the table by processing it diagonally as follows. Start with a_{11}, then count a_{12}, a_{21}. Proceed to a_{31}, a_{22}, a_{13}, etc. Any element of the table will eventually be counted this way, and for any positive integer n we can find a specific element in the table that is counted as n^{th} in the suggested procedure. Thus, we have a bijection between $\mathbb{Z}_+$ and the union of elements of $A_1, A_2, \ldots$.

(c) Let A be an infinite set. We construct its countable subset as follows. Pick an arbitrary element a_1 of A. Then pick an arbitrary element of $A \backslash \{a_i\}$, ..., pick an arbitrary element a_k of $A \backslash \{a_1, \ldots, a_{k-1}\}$ for any $k \geq 3$. This is always possible since the set $A \backslash \{a_1, \ldots, a_{k-1}\}$ is infinite for any positive integer $k \geq 2$.

1.4. Show that each of the following sets satisfies axioms from Definition 1.4 and thus is a linear space.

(a) $\mathbb{R}^n$ with operations of vector addition and scalar multiplication.

(b) $\mathbb{R}^{m \times n}$ with operations of matrix addition and scalar multiplication.

(c) $C[a,b]$–the set of all continuous real-valued functions defined on the interval $[a,b]$, with addition and scalar multiplication defined as

$$(f+g)(x) = f(x) + g(x), x \in [a,b] \text{ for any } f, g \in C[a,b];$$

$$(\alpha f)(x) = \alpha f(x), x \in [a,b] \text{ for each } f \in C[a,b] \text{ and any scalar } \alpha;$$

respectively.

(d) $\mathcal{P}_n$–the set of all polynomials of degree at most n with addition and scalar multiplication defined as

$$(p+q)(x) = p(x) + q(x), \quad \text{for any } p, q \in \mathcal{P}_n;$$

$$(\alpha p)(x) = \alpha p(x), \quad \text{for each } p \in \mathcal{P}_n; \text{and any scalar } \alpha;$$

respectively.

Solution: All 8 properties listed in Definition 1.4 are trivial to verify in each of the four considered cases.

1.5. (**Gram-Schmidt orthogonalization**) Let $p^{(0)}, \ldots, p^{(k-1)} \in \mathbb{R}^n$, where $k \leq n$, be an arbitrary set of linearly independent vectors. Show that the set of vectors $d^{(0)}, \ldots, d^{(k-1)} \in \mathbb{R}^n$ given by

$$\begin{aligned} d^{(0)} &= p^{(0)}; \\ d^{(s)} &= p^{(s)} - \sum_{i=0}^{s-1} \frac{{p^{(s)}}^T d^{(i)}}{{d^{(i)}}^T d^{(i)}} d^{(i)}, \quad s = 1, \ldots, k-1 \end{aligned}$$

is orthogonal.

Solution: We have

$$\begin{aligned} {d^{(0)}}^T d^{(1)} &= {p^{(0)}}^T \left(p^{(1)} - \frac{{p^{(1)}}^T p^{(0)}}{{p^{(0)}}^T p^{(0)}} p^{(0)} \right) \\ &= \frac{{p^{(0)}}^T p^{(1)} ({p^{(0)}}^T p^{(0)}) - {p^{(0)}}^T ({p^{(1)}}^T p^{(0)}) p^{(0)}}{{p^{(0)}}^T p^{(0)}} \\ &= 0. \end{aligned}$$

Assume that ${d^{(j)}}^T d^{(s)} = 0$ for all $s = 1, \ldots, r$, $j = 1, \ldots s-1$, where $r < k-1$. Then for $s = r+1$ and $j = 1, \ldots s-1$ we have

$$\begin{aligned}
{d^{(j)}}^T d^{(r+1)} &= {d^{(j)}}^T \left(p^{(r+1)} - \sum_{i=0}^{r} \frac{{p^{(r+1)}}^T d^{(i)}}{{d^{(i)}}^T d^{(i)}} d^{(i)} \right) \\
&= {d^{(j)}}^T \left(p^{(r+1)} - \frac{{p^{(r+1)}}^T d^{(j)}}{{d^{(j)}}^T d^{(j)}} d^{(j)} \right) \\
&= {d^{(j)}}^T p^{(r+1)} - {p^{(r+1)}}^T d^{(j)} \\
&= 0.
\end{aligned}$$

Hence, the statement is correct by induction.

1.6. Show that for $p = 1, 2, \infty$ the norm $\|\cdot\|_p$ is *compatible*, that is for any $A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n$:

$$\|Ax\|_p \le \|A\|_p \|x\|_p.$$

Solution: By the definition of $\|A\|_p$ we have:

$$\|A\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} \ge \frac{\|Ax\|_p}{\|x\|_p} \quad \text{for any } x \in \mathbb{R}^n \backslash \{0\}.$$

Thus, $\|Ax\|_p \le \|A\|_p \|x\|_p$.

1.7. Let

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

be two vectors in $\mathbb{R}^n$.

(a) Compute the matrix

$$M = ab^T = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} [b_1, \ldots, b_n].$$

(b) How many additions and multiplications are needed to compute M?

(c) What is the rank of M?

Solution:

(a) By definition of the product of two matrices,

$$M = ab^T = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} [b_1, \ldots, b_n] = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & \ldots & a_n b_n \end{bmatrix}.$$

(b) To compute M in (a), we have done n^2 multiplications and no additions.

(c) Without loss of generality, assume that $a_1 \neq 0$. Then the i-th row R_i, $i = 2, \ldots, n$, of M can be obtained from the first row R_1 as follows: $R_i = \frac{a_i}{a_1} R_1$. Therefore, the rank of M is 1.

1.8. Show that a square matrix A is orthogonal if and only if both the columns and rows of A form sets of orthonormal vectors.

Solution: Denote by a_i the i^{th} column of A of size $n \times n$. Then a_i^T is the i^{th} row of A^T. Since A is orthonormal have: $A^T A = I_n$, where I_n is the $n \times n$ identity matrix. But this is the case if and only if

$$a_i^T \times a_j = \begin{cases} 1, \text{ if } i = j \\ 0, \text{otherwise,} \end{cases}$$

which is exactly the definition of a set of orthonormal vectors. Since $A^T = A^{-1}$, we have $AA^T = I_n$ and the same proof can be used to show that the rows of A are also orthonormal.

1.9. A company manufactures three different products using the same machine. The sell price, production cost, and machine time required to produce a unit of each product are given in the following table.

	Product 1	Product 2	Product 3
Sell price	\$30	\$35	\$50
Production cost	\$18	\$22	\$30
Machine time	20 min	25 min	35 min

The table below represents a two-week plan for the number of units of each product to be produced.

	Week 1	Week 2
Product 1	120	130
Product 2	100	80
Product 3	50	60

Use a matrix product to compute, for each week, the revenue received from selling all items manufactured in a given week, the total production cost for each week, and the total machine time spent each week. Present your answers in a table.

Solution: We have

$$\begin{bmatrix} 30 & 35 & 50 \\ 18 & 22 & 30 \\ 20 & 25 & 35 \end{bmatrix} \begin{bmatrix} 120 & 130 \\ 100 & 80 \\ 50 & 60 \end{bmatrix} = \begin{bmatrix} 9600 & 9700 \\ 5860 & 5900 \\ 6650 & 6700 \end{bmatrix}.$$

Putting the results in a table, we obtain:

	Week 1	Week 2
Revenue	9600	9700
Production cost	5860	5900
Machine time	6650	6700

1.10. Given matrices

$$A = \begin{bmatrix} 1 & 2 & 6 \\ -2 & 7 & -6 \\ 2 & 10 & 5 \\ 0 & 4 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ 1 & -3 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

(a) find the transpose of each matrix;

(b) consider all possible pairs $\{(A,A),(A,B),(B,A),\ldots,(C,C)\}$ and compute a product of each pair for which multiplication is a feasible operation.

Solution:

(a)

$$A^T = \begin{bmatrix} 1 & -2 & 2 & 0 \\ 2 & 7 & 10 & 4 \\ 6 & -6 & 5 & 8 \end{bmatrix}, \quad B^T = \begin{bmatrix} 2 & 1 & 0 \\ -1 & -3 & 2 \end{bmatrix}, \quad \text{and} \quad C^T = C.$$

(b)

$$AB = \begin{bmatrix} 4 & 5 \\ 3 & -31 \\ 14 & -22 \\ 4 & 4 \end{bmatrix}; \; AC = A; \; CB = B; \; CC = C.$$

1.11. For the matrices

$$A = \begin{bmatrix} -1 & 8 & 4 \\ 2 & -3 & -6 \\ 0 & 3 & 7 \end{bmatrix}; \quad B = \begin{bmatrix} -4 & 9 & 2 \\ 3 & -5 & 4 \\ 8 & 1 & -6 \end{bmatrix},$$

find (a) $A - 2B$, (b) AB, (c) BA.

Solution:

$$\text{(a) } A - 2B = \begin{bmatrix} 7 & -10 & 0 \\ -4 & 7 & -14 \\ -16 & 1 & 19 \end{bmatrix}; \text{ (b) } AB = \begin{bmatrix} 60 & -45 & 6 \\ -65 & 27 & 28 \\ 65 & -8 & -30 \end{bmatrix};$$

$$\text{(c) } BA = \begin{bmatrix} 22 & -53 & -56 \\ -13 & 51 & 70 \\ -6 & 43 & -16 \end{bmatrix}.$$

1.12. Compute the p-norm for $p = 1, 2, \infty$ of the matrices A and B in Exercise 1.11.

Solution: $\|A\|_1 = 17, \; \|A\|_2 = 12.882, \; \|A\|_\infty = 13;$
$\|B\|_1 = 15, \; \|B\|_2 = 12.168, \; \|B\|_\infty = 15.$

1.13. Find the quadratic Taylor's approximation of the following functions:

(a) $f(x) = x_1^4 + x_2^4 - 4x_1x_2 + x_1^2 - 2x_2^2$ at $\bar{x} = [1,1]^T$;

(b) $f(x) = \exp(x_1^2 + x_2^2)$ at $\bar{x} = [0,0]^T$;

(c) $f(x) = \frac{1}{1+x_1^2+x_2^2}$ at $\bar{x} = [0,0]^T$.

Solution: The quadratic Taylor's approximation is given by

$$f(x) \approx f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}).$$

(a) $f(\bar{x}) = 1 + 1 - 4 + 1 - 2 = -3.$

$\nabla f(x) = [4x_1^3 - 4x_2 + 2x_1, 4x_2^3 - 4x_1 - 4x_2]^T \quad \Rightarrow \quad \nabla f(\bar{x}) = [2, -4]^T.$

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 + 2 & -4 \\ -4 & 12x_2^2 - 4 \end{bmatrix} \quad \Rightarrow \quad \nabla^2 f(\bar{x}) = \begin{bmatrix} 14 & -4 \\ -4 & 8 \end{bmatrix}.$$

Hence, we have:

$$\begin{aligned} f(x) \quad &\approx \quad -3 + 2(x_1 - 1) - 4(x_2 - 1) + 7(x_1 - 1)^2 \\ &\quad\quad -4(x_1 - 1)(x_2 - 1) + 4(x_2 - 1)^2 \\ &= \quad 7x_1^2 - 4x_1x_2 + 4x_2^2 - 8x_1 - 8x_2 + 6. \end{aligned}$$

(b) $f(\bar{x}) = \exp(0 + 0) = 1.$

$\nabla f(x) = [2x_1 \exp(x_1^2 + x_2^2), 2x_2 \exp(x_1^2 + x_2^2)]^T \quad \Rightarrow \quad \nabla f(\bar{x}) = [0, 0]^T.$

$$\nabla^2 f(x) = 2\exp(x_1^2 + x_2^2) \begin{bmatrix} 1 + 2x_1 & 2x_1x_2 \\ 2x_1x_2 & 1 + 2x_2 \end{bmatrix}$$

$$\Rightarrow \quad \nabla^2 f(\bar{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Thus,

$$f(x) \approx x_1^2 + x_2^2 + 1.$$

(c) $f(\bar{x}) = 1.$

$$\nabla f(x) = [-\frac{2x_1}{(1 + x_1^2 + x_2^2)^2}, -\frac{2x_2}{(1 + x_1^2 + x_2^2)^2}]^T \quad \Rightarrow \quad \nabla f(\bar{x}) = [0, 0]^T.$$

$$\nabla^2 f(x) = \begin{bmatrix} -\frac{2(1+x_1^2+x_2^2)^2 - 8(1+x_1^2+x_2^2)x_1^2}{(1+x_1^2+x_2^2)^4} & \frac{4(1+x_1^2+x_2^2)x_1x_2}{(1+x_1^2+x_2^2)^4} \\ \frac{4(1+x_1^2+x_2^2)x_1x_2}{(1+x_1^2+x_2^2)^4} & -\frac{2(1+x_1^2+x_2^2)^2 - 8(1+x_1^2+x_2^2)x_2^2}{(1+x_1^2+x_2^2)^4} \end{bmatrix}$$

$$\Rightarrow \quad \nabla^2 f(\bar{x}) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Thus,

$$f(x) \approx -x_1^2 - x_2^2 + 1.$$