Math164 – Homework for "Stephen G. Nash and Ariela Sofer, *Linear and Nonlinear Programming*, McGraw-Hill (1996)

Andrea Brose, Winter '05

Here is the *hopefully* complete list of homework assignments you will be doing in the course of this quarter. The exercise numbers in each section refer to the exercises in the text. If the assignment on the assignment page says you should do section $n.m$, where $n \in \{1, 2, 3, 4, 5, 6, 10, 14, A, B\}$, $m \in \{1, 2, 3, 4, 5, 6, 7\}$, it means you should do all the problems following $n.m$ *including* the additional so called AP problems corresponding to the section, unless otherwise stated.
Example: WA5, due 13/32/0000: sections 6.1, 6.2 means that the following problems are assigned: 1, 2, 3, 4, 7 in section 6.1 of the text, AP(vii) and AP(viii) and 1, 3, 4, 5, 7(a)(d), 8(a)(c), 9 in section 6.2.

Section Exercises

1.1-1.5 For problems AP(i)–(iii), write out the programming problem and classify the problem according to

- linear program
- nonlinear unconstrained program
- nonlinear problem with linear constraint(s)
- nonlinear problem with nonlinear constraint(s)

AP(i) A cheese shop has 10 kg of seasonal fruit mix and 30 kg of an expensive cheese with with which it will make two cheese spreads, deluxe and regular. Each kg of the deluxe spread consists of 200 g of the fruit mix and 800 g of the expensive cheese, while each kg of the regular spread consists of 200 g of the fruit mix, 300 g of the expensive cheese and 500 g of the filler cheese, which is cheap and plentiful in supply. From past pricing policies, the shop has found that the demand for each spread depends on its price as follows:

$$D_1 = 190 - 25P_1 \qquad \text{and} \qquad D_2 = 250 - 50P_2$$

where D_i, denotes demand in kg and P_i denotes price (in € per kg) for $i = 1$ for the deluxe and $i = 2$ the regular spread respectively. How many kg of each spread should the cheese shop prepare, and what prices should it establish, if it wishes to maximize income and be left with no inventory of either spread at the end of the season.

AP(ii) The Dörfli baker traditionally makes his bread from a combination of spelt and rye flour. Per 100g grain spelt contains 8.7 grams of protein while rye contains 11.7 grams. Spelt costs \$1.52 per kg and rye \$1.20 per kg. How much of each kind of grain should the baker use in each bread made of 1kg flour if he wants to minimize the cost while keeping the protein content to no less than 110g. (Note that the other ingredients being water, salt, sugar and yeast do not contribute any more protein to the bread.)

AP(iii) A major oil company wants to build a refinery that will be supplied from three port cities. Port B is located 300 km East and 400 km north of port A, while port C is 400 km east and 100 km south of port B. Determine the location of the refinery so that the total amount of pipe required to connect the refinery to the ports is minimized.

2.2 1-4, 6

2.3 1, 3, 4, 6[1], 7-11, 13

2.4 1, 2, 3

3.1 1, 2, 3(a)(b)(c)(d), 5

4.1 1(b)(c)(f)(g), 2

4.2 1, 2, 3, problem 8 from section 2.2 on page 20, and

In class we showed that if we add a slack variable to a "$\leq$" constraint of a linear programming problem LP1 that the new linear programming problem LP2 is 'equivalent' to the original one, in the sense that any optimal solution of LP1 is also an optimal solution of LP2 and vice versa.

AP(iv) Show that the same is true when we subtract a surplus variable in a "$\geq$" constraint.

AP(v) Show that the same is true if the objective in LP1 is to be maximized and hence transformed to be minimized in LP2.

4.3 1 (You may stop after you found some basic solutions, if you think you understood the concepts, as there are 10 of those.), 2, 3(a), 4, 5, 6(a)(b), 9, 12 (Hint: Construct $\boldsymbol{y}$ and $\boldsymbol{z}$.), 14 (Hint: try to do the problem for $n = 2$ first. Once you understood this, generalize your answer for every n.)

AP(vi) Consider the matrix

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

and the vector

$$\begin{bmatrix} 2 \\ 1 \\ c \end{bmatrix},$$

where $c \in \mathbb{R}$ is a constant.

Choose c such that

(a) the system $A\boldsymbol{x} = \boldsymbol{b}$ has no solutions,

(b) the system $A\boldsymbol{x} = \boldsymbol{b}$ has so called redundant equations.

4.4 1 (Hint: use theorem 4.1.), 4, 5, 6, 7

[1] See appendix for a definition of a convex combination of functions.

5.2 2(b)(c)(d)(add $x_3 \geq 0$)(e), 3 (Solve graphically and also by the simplex method.), 5, 8[2]

6.1 1, 2, 3, 4, 7 and two additional problems for section 6.1:

The idea of the two exercises below is, to understand and convince yourself of the 'recipe' that is given on the top of page 148. After having done these exercises, try to do example 6.2 on page 148 (without looking at the solution of course). Note that there are two typos: in the dual, the period in the last row should be a comma, but more importantly, in the paragraph following the dual explaining the derivation "minimization" in the paragraph's second to last line should be "maximization".

AP(vii) First consider a primal minimization problem with a mix of "≤", "≥" and "=" constraints:

$$\left.\begin{array}{ll} \text{minimize} & z = 7x_1 - 8x_2 \\ \text{subject to} & x_1 + x_2 \geq 3 \\ & 4x_1 - 2x_2 \leq 6 \\ & -x_1 + x_2 = 1 \\ & \boldsymbol{x} \geq 0 \end{array}\right\} \text{(P1)}$$

(a) Find the canonical form of the above primal (P1).
(b) Find the dual of (P1) in CF.
(c) Try to explain the pattern, i.e. what are the differing impacts of the various constraints on the dual. 'Hint': see pages 146-147 in the text.

AP(viii) Now consider a primal minimization problem with a mix of non-negative, non-positive and free (unrestricted) variables:

$$\left.\begin{array}{ll} \text{minimize} & z = 3x_1 - 2x_2 + 4x_3 \\ \text{subject to} & x_1 + x_2 + x_3 \geq 3 \\ & 2x_1 + 4x_2 - x_3 \geq 1 \\ & x_1 \geq 0, x_2 \leq 0, x_3 \text{ free} \end{array}\right\} \text{(P2)}$$

(a) Find the canonical form of the above primal (P2).
(b) Find the dual of (P2).
(c) Try to explain the pattern, i.e. what are the differing impacts of the various variables on the dual. 'Hint': see page 147 in the text.

Note that these two examples, as well as the table on page 148 gives you only a small generalization.

6.2 1, 3[3], 4, 5, 7(a)(Hint: use the ideas of the simplex method)(d), 8(a)(c), 9

6.2.1 12, 13, 14, 15

[2] ".. feasible problem" means $S \neq \emptyset$. Also a hint: just as with problem 14 in 4.3, try to do the problem for $n = 2$ first. Once you understood this, generalize your answer for every n.

[3] See appendix for an explanation of "... the primal is unbounded ..." .

A.6

AP(ix) Check whether $A = \begin{bmatrix} 9 & -1 & 2 \\ -1 & 7 & -3 \\ 2 & -3 & 3 \end{bmatrix}$ is positive or negative definite.

AP(x) For which values of p and q is $\begin{bmatrix} p & q & q \\ q & p & q \\ q & q & p \end{bmatrix}$ positive definite?

B.4

AP(xi) Find the gradient of $f(x_1, x_2, x_3) = \sin(x_1 x_2^2 - e^{x_3}) + x_1^{x_2}$.

AP(xii) Find the Hessian of $f(x_1, x_2, x_3) = x_1^2 + x_1 x_2 + x_1 x_3^2$.

AP(xiii) find the Jacobian of $f(x_1, x_2, x_3) = \begin{bmatrix} x_1 x_3^2 + \sin x_1 x_2^2 \\ x_1^2 + x_2^2 + x_3^2 \end{bmatrix}$.

2.3.1 14(b)(d)(g), 15, 16

2.6 1, 2(a)(b), 4, 5, 6 (Remark and definition: if $\boldsymbol{p}$ is such that $\boldsymbol{p}^T \nabla f(\boldsymbol{x}) < 0$ then $\boldsymbol{p}$ is called feasible *descent* direction at $\boldsymbol{x}$.)

2.7 1 (Choose $\boldsymbol{x}^{(0)} = 1$ and calculate $\boldsymbol{x}^{(2)}$ and $f(\boldsymbol{x}^{(2)})$ to four decimal places.), 4 (First task only!), 10 (Choose $\boldsymbol{x}^{(0)} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ and calculate $\boldsymbol{x}^{(1)}$ and $f(\boldsymbol{x}^{(1)})$ to four decimal places.)

10.2 1(a), 3, 4 (Omit the *second-order necessary condition* in part (a).), 6, 7, 10(a) (Hint: See B5.), 12, 17

For the first two problems, an optimum may not exist, in which case you should argue why.

AP(xiv) Maximize $f(x_1, x_2, x_3) = x_1 x_2 - x_1 + x_3^3 - 3x_3$.

AP(xv) Minimize $f(x_1, x_2) = (x_1 - \sqrt{5})^2 + (x_2 - 5)^2 + 10$.

AP(xvi) Minimize $f(x_1, x_2) = \sin x_1 x_2 - \cos(x_1 - x_2)$. Hint: Once you found the gradient, you will notice that you cannot explicitely solve $\nabla f = \mathcal{O}$, and hence won't find all stationary points. Proceed as follows:

- Find a lower bound for f.
- Find a stationary point by 'looking' at the gradient long enough and justify that it must be a global minimum.

AP(xvii) Prove the **S**econd **O**rder **S**ufficient **C**ondition for functions of several variables, i.e. if

$$\begin{aligned} f \,:\, U \subset \mathbb{R}^n &\to \mathbb{R} \\ (x_1, \ldots, x_n) &\mapsto f(x_1, \ldots, x_n) \end{aligned}$$

is such that

$$\nabla f(\boldsymbol{x}^*) = \mathcal{O} \qquad \text{and} \qquad \nabla^2 f(\boldsymbol{x}^*) \text{is positive definite,}$$

then $\boldsymbol{x}^*$ is a local minimizer, in fact a *strict* local minimizer of f.

10.3 2 (Choose $\boldsymbol{x}^{(0)} = 1$, calculate $\boldsymbol{x}^{(3)}$, calculate $f'(\boldsymbol{x}^{(3)})$ and verify that $f''(\boldsymbol{x}^{(3)}) > 0$.), 3 (Calculate $\boldsymbol{x}^{(1)}$ and verify that the Hessian is positive definite at $\boldsymbol{x}^{(1)}$.), 6 (Compare your answer with problem 4 in 2.7) and

AP(xviii) Use Newton's method to find an approximate solution to AP(xi): Minimize $f(x_1, x_2) = \sin x_1 x_2 - \cos(x_1 - x_2)$ with starting point $\boldsymbol{x}^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} = \begin{bmatrix} 0.046 \\ 0.081 \end{bmatrix}$.

AP(xix) In this exercise you show that the Jacobian of the Gradient is the Hessian in the case of a function of two variables only. If

$$\begin{aligned} f \,:\, U \subset \mathbb{R}^2 &\to \mathbb{R} \\ (x_1, x_2) &\mapsto f(x_1, x_2) \end{aligned}$$

is such that $f \in C^2$, show that

$$\boldsymbol{\nabla}(\boldsymbol{\nabla} f(\boldsymbol{x}))^T = \boldsymbol{\nabla}^2 f(\boldsymbol{x}) \,.$$

Note that we have the transpose on the left hand side, yet there is no transpose on the right hand side. Do not forget to argue why that is the case.

3.2 1(a)(b)(c) (Find $\boldsymbol{x}^{(1)}$ only.), 2, 3, 5 (Assume A has full row rank.)

B.7 For the following problems, assume differentiability for all functions as needed.

AP(xx) Let $A \in M_{m\times n}(\mathbb{R})$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^m$. For the parts below, find $\boldsymbol{\nabla} f$.

(a) $f(\boldsymbol{x}) = \boldsymbol{y}^T A \boldsymbol{x}$

(b) $f(\boldsymbol{x}) = \boldsymbol{x}^T A^T \boldsymbol{y}$

(c) $f(\boldsymbol{x}) = \boldsymbol{y}^T \boldsymbol{x}$

(d) $f(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$ (here $m = n$)

(e) $f(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{x}$

AP(xxi) Find the derivatives of the following functions:

(a) $H(t) = h(h(t))$, where $h \,:\, \mathbb{R} \to \mathbb{R}$.

(b) $G(t) = g(t, -t)$, where $g \,:\, \mathbb{R}^2 \to \mathbb{R}$

(c) $F(t) = f(t, f(t, -t))$, where $f \,:\, \mathbb{R}^2 \to \mathbb{R}$.

AP(xxii) (a) $\boldsymbol{x}(\boldsymbol{t}) = \begin{bmatrix} \boldsymbol{t} \\ \boldsymbol{y}(\boldsymbol{t}) \end{bmatrix}$, where $\boldsymbol{y} \,:\, \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \,:\, \mathbb{R}^n \to \mathbb{R}^{n+m}$ and $\boldsymbol{t} \in \mathbb{R}^n$. Find $\boldsymbol{\nabla x}$.

(b) $\boldsymbol{f}(\boldsymbol{t}) = \boldsymbol{g}(\boldsymbol{x}(\boldsymbol{t}))$, where $\boldsymbol{x}$ is as in the part (a) and hence $\boldsymbol{g} \,:\, \mathbb{R}^{n+m} \to \mathbb{R}^r$. Find $\boldsymbol{\nabla f}$.

14.2 1, 2 (Rather than "solving" this problem, find local solutions only.), 4 (Hint: use Lagrange Multipliers), 5[4] (Hint: use the ideas from p. 429, 2nd ¶) ,

Following is a warm up for problem 6 in section 14.2. You can 'solve' this problem by simply applying theorem 2.1 (Global Solutions for Convex Programs), but I want you to solve it differently, to warm up to problem 6 in 14.2:

[4] Correct the hypothesis "... *if the problem has a feasible solution* ... " to "... *if the problem has a feasible point* ... " i.e. $S \neq \emptyset$.

AP(xxiii) Consider

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x} \\ \text{subject to} \quad & A\boldsymbol{x} = \boldsymbol{b}\,, \end{aligned}$$

where Q is symmetric, positive definite, $n \times n$ matrix, $\boldsymbol{x} \in \mathbb{R}^n$, A is regular $m \times n$ matrix and $\boldsymbol{b} \in \mathbb{R}^m$.
Show that any local minimizer is also a global minimizer by studying the following:
Let $\boldsymbol{x}^*$ be a local minimizer. Express each $\boldsymbol{x} \in S = \{\boldsymbol{x} \,:\, A\boldsymbol{x} = \boldsymbol{b}\}$ as

$$\boldsymbol{x} = \boldsymbol{x}^* + Z\boldsymbol{v}\,,$$

where Z is $n \times (n-m)$ basis null space matrix of A and $\boldsymbol{v} \in \mathbb{R}^{n-m}$.

6 (As in the warm up to this problem, assume that Q is symmetric, positive definite matrix.), 7 (Hint: see B.7 on page 653 in the text.), 9 and

AP(xxiv) In class we used the following fact:
Let $A \in M_{m\times n}(\mathbb{R})$, consider

$$S = \{\boldsymbol{x} \,:\, A\boldsymbol{x} = \boldsymbol{b}\}\,,$$

and let $\bar{\boldsymbol{x}} \in S$, then for all $\boldsymbol{x} \in S$ there exists a $\boldsymbol{p} \in \mathcal{N}(A)$ such that

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{p}$$

Prove the statement.

14.3 1, 2, 3

14.4 1, 2 (Rather than "solving" this problem, find local solutions to it.), 3, 5 (In part (a) omit SONC.), 6

14.5 3(a)(Hint: See B5.)(b), 4, 5

AP(xxv) Maximize $f(x_1, x_2) = 2x_1 + x_1x_2 + 3x_2$ subject to $x_1^2 + x_2 = 3$.

AP(xxvi) Find local solutions to minimize $f(x_1, x_2) = x_1^2 x_2$ subject to $x_1^2 + 2x_2^2 = 6$

AP(xxvii) Find local solutions to minimize $f(x_1, x_2, x_3) = x_1 + 2x_2$ subject to $x_1 + x_2 + x_3 = 1$ and $x_2^2 + x_3^2 = 4$

AP(xxviii) Find local solutions to minimize $f(x_1, x_2, x_3) = x_1x_2 + x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$.

AP(xxix) Find local solutions to minimize $f(x_1, x_2) = (x_1-2)^2 + (x_2-1)^2$ subject to $x_1 - 2x_2 = -1$ and $x_1^2 + 4x_2^2 \leq 4$.

14.7 1, 2[5], 3

[5]..."feasible arc', is a feasible curve.

Appendix

2.3 Problem 6

What is a *convex combination of convex functions?* On page 22 we learned what a *convex combination of points* is. Similarly we can define the convex combination of functions (or simply think of functions as points, if that is not confusing you). Here it goes, nevertheless:

DEFINITION: Given k functions $f_1, f_2, \ldots, f_k$ (not necessarily convex themselves), all defined on the same set S, f is said to be a *convex combination of the functions* $\{f_i\}_{i=1}^k$ if

$$f(\boldsymbol{x}) = \sum_{i=1}^{k} \alpha_i f_i(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k \geq 0$ are non-negative scalars, such that $\sum_{i=1}^k \alpha_i = 1$.

EXAMPLE: Let $S = \mathbb{R}$, $k = 3$ and $f_1(x) = x^2$, $f_2(x) = \sin x$, and $f_3(x) = e^x$, then

$$f(x) = \frac{1}{2}x^2 + \frac{1}{4}\sin x + \frac{1}{4}e^x$$

is a convex combination of the $\{f_i\}_{i=1}^3$, with coefficients $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \alpha_3 = \frac{1}{4}$.

6.2 Problem 3

Concerning corollary 6.1 on page 151 of the textbook: What does

> ... the primal is unbounded ...

mean? The unboundedness refers to the *objective* being unbounded, in specific, as the primal is (when in canonical form) a minimization problem, it means the primal is unbounded *below* i.e. $z \to -\infty$ for feasible $\boldsymbol{x}$.

Similarly for the second statement of Corollary 6.1:

> ... the primal is unbounded ...

means that the objective that is to be maximized (for the canonical form of the dual) is unbounded above, i.e. $z \to \infty$ for feasible $\boldsymbol{y}$.