

U. S. AIR FORCE

PROJECT RAND

RESEARCH MEMORANDUM

SOME RESULTS IN NONLINEAR PROGRAMMING
PART II

R. M. Thrall

RM-935

5 September 1952

Assigned to _____

This is a working paper. It may be expanded,
modified, or withdrawn at any time.

The **RAND** *Corporation*

1700 MAIN ST. • SANTA MONICA • CALIFORNIA

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 05 SEP 1952		2. REPORT TYPE		3. DATES COVERED 00-00-1952 to 00-00-1952	
4. TITLE AND SUBTITLE Some Results in Nonlinear Programming Part II				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Rand Corporation, Project Air Force, 1776 Main Street, PO Box 2138, Santa Monica, CA, 90407-2138				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 16	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Summary: The solution of a minimization problem is obtained in the vector case; its properties are studied and applied to a maximization problem.

SOME RESULTS IN NONLINEAR PROGRAMMING, PART II

R. M. Thrall

§1. Introduction.

In the present note we consider a class of maximization and minimization problems which fall under the general heading of nonlinear programming. The functions which enter are assumed to satisfy enough differentiability conditions so that methods of calculus can be applied and so that existence of at least one solution is trivial. This note is an extension and generalization of RM-909. J. Danskin has considered a similar problem for the functional case in RM-618 and this has been generalized by J. Danskin and H. Kahn in a forthcoming memorandum.

§2. Notation and Geometric Preliminaries.

Let V_n be the space of all real vectors $v = (v_1, \dots, v_n)$. For any vector v we define v_0 to be the sum of the components of v . A vector v is said to be positive, written $v > 0$ ($0 = (0, \dots, 0)$) if each component is positive. We write

$v \geq 0$ if each component is non negative and write $v \geq 0$ if $v \geq 0$ but $v \neq 0$. If $v_0 = 1$ and $v \geq 0$ we call v a probability vector.

For any $x_0 > 0$, we denote by $\Gamma = \Gamma(x_0)$ the set of all vectors x in V_n for which $x \geq 0$ and $x_1 + \dots + x_n = x_0$. Let E be any subset of $\{1, \dots, n\}$. We denote by Γ_E the set of all vectors x in Γ for which $x_i \neq 0$, i in E and $x_i = 0$, i not in E . We call the subsets Γ_E the faces of Γ . Clearly, Γ is partitioned by its faces.

§ 3. Functions of Type A.

Let $q(t)$ be a real valued function defined for $t \geq 0$ and with the following properties:

- (A) $\left\{ \begin{array}{ll} \text{(i)} & q(t) \text{ monotone decreasing} \\ \text{(ii)} & q''(t) > 0 \text{ for all } t > 0 \\ \text{(iii)} & \lim_{t \rightarrow \infty} q(t) = 0 \\ \text{(iv)} & q(0) = 1 \end{array} \right.$

Actually, properties (iii) and (iv) could be replaced by the single assumption $\lim_{t \rightarrow \infty} q(t) > -\infty$, but then by a normalization we could regain (iii) and (iv). A function which satisfies conditions A is said to be of type A. We observe that one consequence of (i) and (ii) is that $q'(t)$ is a monotone increasing function with negative values.

§ 4. The Minimization Problem.

Let P be a positive probability vector, let $q_1(t), \dots, q_n(t)$ be functions of type A and consider minimization of the function

$$(1) \quad f(x, p) = \sum_{i=1}^n p_i q_i(x_i)$$

for vectors x in $\Gamma(x_0)$. Since $f(x, p)$ is continuous in x and since $\Gamma(x_0)$ is compact the minimum exists; we denote it by

$$(2) \quad g(x_0) = \min_{x \in \Gamma(x_0)} f(x, p) .$$

Let $\bar{x}$ be a minimizing vector and suppose that $\bar{x} \in \Gamma_E$. Let $j_0 \in E$ and write

$$(3) \quad x_{j_0} = 1 - \sum_{j \neq j_0} x_j .$$

For $j \neq j_0$, we have

$$(4) \quad \frac{\partial f(x, p)}{\partial x_j} = p_j q'_j(x_j) - p_{j_0} q'_{j_0}(x_{j_0}) .$$

In particular, for $x = \bar{x}$ we must have

$$(5) \quad \left[\frac{\partial f}{\partial x_j} \right]_{x=\bar{x}} \geq 0 \quad j \neq j_0$$

with equality holding for $j \in E$.

From (4) this gives

$$(6) \quad p_j q_j'(\bar{x}_j) = p_{j_0} q_{j_0}'(\bar{x}_{j_0}) = \mu_E \quad (j \in E)$$

and

$$(7) \quad p_j q_j'(0) \geq \mu_E \quad (j \notin E).$$

Let

$$(8) \quad \rho_j = p_j q_j'(0) \quad (j=1, \dots, n).$$

Now since $q_j'(t)$ is a monotone increasing function, it follows from (6), (7), and (8) that

$$(9) \quad \begin{cases} \mu_E > \rho_j & (j \in E) \\ \mu_E \leq \rho_j & (j \notin E). \end{cases}$$

Next, we arrange the indices so that

$$(10) \quad \rho_1 \geq \rho_2 \geq \dots \geq \rho_n,$$

and define $\rho_0 = 0$, $\rho_{n+1} = -\infty$.

It then follows from (9) that E must be one of the sets

$$(11) \quad E_h = \{h, \dots, n\} \quad (h=1, \dots, n).$$

This has already cut down the possible locations Γ_E for minimizing vectors from 2^n to n . We next show that there is only one h for which E_h satisfies (9). We write μ_h for

μ_{E_h} and Γ_h for Γ_{E_h} . Then (9) is replaced by

$$(12) \quad \rho_{h-1} \geq \mu_h > \rho_h$$

where $1 \leq h \leq n+1$.

Let $r_j(w)$ be the inverse of $q_j'(t)$, i.e.,

$$(13) \quad r_j(q_j'(t)) = t, \quad q_j'(r_j(w)) = w \quad (j=1, \dots, n).$$

The domain of r_j is $\rho_j/p_j \leq w < 0$ and its range is $0 \leq r_j(w) < \infty$.

If x is a maximizing vector which is in Γ_h we have as a consequence of (6) that

$$(14) \quad \begin{cases} \bar{x}_j = 0 & (j < h) \\ \bar{x}_j = r_j(\mu_h/p_j) & (j \geq h) \end{cases}$$

where $\mu_h = \mu_h(x_0)$ is defined as a function of x_0 by the equation

$$(15) \quad x_0 = \sum_{j=h}^n r_j(\mu_h/p_j).$$

Since each $q_j(t)$ is of type A, the inverse functions $r_j(w)$ are monotone increasing. Hence equation (15) has a unique solution μ_h which is a monotone increasing function of x_0 with domain

$$x_0 \geq y_h = \sum_{j=h}^n r_j(\rho_h/p_j)$$

and range $\rho_h \leq \mu_h < 0$. Moreover, we have for $x_0 \geq y_h$ that

$$(16) \quad \mu_h(x_0) \leq \mu_{h+1}(x_0)$$

and equality holds in (16) if and only if $r_j(\mu_j(x_0)/p_j) = 0$; i.e., if and only if $\mu_h(x_0) = \rho_h$. But now it follows from (15) that

$$(17) \quad \begin{cases} \mu_h(y_h) = \mu_{h+1}(y_h) = \rho_h & (h=1, \dots, n-1) \\ \mu_n(y_n) = \rho_n. \end{cases}$$

We set $y_0 = \infty$ and then have $0 = y_n \leq y_{n-1} \leq \dots \leq y_1 \leq y_0$.

Formula (17) provides the clue for the determination of an h which satisfies (12). Indeed, (12) holds if and only if h satisfies the condition

$$(18) \quad y_h < x_0 \leq y_{h-1}.$$

Since (18) has only one solution we have established uniqueness for the minimizing vector. We summarize these results in the following theorem.

Theorem 1. Let $q_1(t), \dots, q_n(t)$ be functions of type A,
let p be a positive probability vector, and let x_0 be a
positive real number. Then the function

$$f(x, p) = \sum_{i=1}^n p_i q_i(x_i)$$

where x has domain $\prod(x_0)$ has a unique minimizing vector $\bar{x}$

given by (14) where h is determined by (18) and μ_h by (15); the minimum value is

$$(19) \quad f(\bar{x}, p) = \sum_{j=1}^{h-1} p_j + \sum_{j=h}^n p_j q_j(r_j(\mu_h/p_j)).$$

The minimum value of $f(x, p)$ can be considered as a function of x_0 , viz

$$(20) \quad g(x_0) = f(\bar{x}, p).$$

For this function we have the following theorem:

Theorem 2. The minimum $g(x_0)$ of $f(x, p)$ is a differentiable monotone decreasing function of x_0 in the interval $x_0 > 0$; moreover, $g'(x_0)$ is continuous for all $x_0 > 0$ and

$$(21) \quad \lim_{x_0 \rightarrow 0^+} g'(x_0) = \rho_n.$$

The second derivative $g''(x_0)$ is a continuous, positive function of x_0 except possibly for the values $y_1, \dots, y_{n-1}$; $g''(x_0)$ is continuous at y_h if and only if $\lim_{t \rightarrow 0^+} q_j''(t) = \infty$, but in any case the one-sided limits exist and are positive. Moreover, $\lim_{x_0 \rightarrow 0^+} g''(x_0)$ exists if and only if $\lim_{t \rightarrow 0} q_n''(t)$ exists.

We first establish the continuity of $g(x_0)$. We have

$$\lim_{x_0 \rightarrow y_h^+} g(x_0) = \sum_{j=1}^{h-1} p_j + \sum_{j=h}^n p_j q_j(r_j(\rho_t/p_j))$$

and

$$\lim_{x_0 \rightarrow y_h^-} g(x_0) = \sum_{j=1}^h p_j + \sum_{j=h+1}^n p_j q_j(r_j(\rho_t/p_j)).$$

The difference is

$$p_h - p_h q_h(r_h(\rho_h/p_h))$$

and this is zero since $r_h(\rho_h/p_h) = 0$ and $q_h(0) = 1$.

Next, for $y_h < x_0 \leq y_{h-1}$ we have

$$\begin{aligned} (22) \quad g'(x_0) &= \sum_{j=h}^n p_j q'_j(r_j(\mu_h/p_j)) \cdot r'_j(\mu_h/p_j) \cdot \frac{1}{p_j} \cdot \frac{d\mu_h}{dx_0} \\ &= \sum_{j=h}^n p_j \frac{\mu_h}{p_j} \cdot \frac{1}{p_j} r'_j(\mu_h/p_j) \frac{d\mu_h}{dx_0}. \end{aligned}$$

Differentiating (15) we get

$$(23) \quad 1 = \sum_{j=h}^n \frac{1}{p_j} r'_j(\mu_h/p_j) \cdot \frac{d\mu_h}{dx_0}.$$

Now from (22) and (23) we get

$$(24) \quad g'(x_0) = \mu_h \quad (y_h < x_0 \leq y_{h-1}).$$

$$\text{Now } \lim_{x_0 \rightarrow y_t^+} \mu_h = \lim_{x_0 \rightarrow y_t^-} \mu_h = \rho_h \text{ and } \lim_{x_0 \rightarrow y_n^+} \mu_h = \rho_n.$$

This establishes all of the statements about $g'(x_0)$.

Finally, for $y_h < x_0 \leq y_{h-1}$ we have

$$(25) \quad g''(x_0) = \frac{d\mu_h}{dx_0} = 1 / \sum_{j=h}^n \frac{1}{p_j} r'_j(\mu_h/p_j).$$

Now, differentiating (13) we get

$$(26) \quad r'_j(q'_j(t) q''_j(t)) = 1$$

and then since $\mu_h = p_j q'_j(\bar{x}_j)$ ($j=h, \dots, n$) and since $q_j(t)$ is of type A, we have

$$(27) \quad r'_j(\mu_h/p_j) = 1/q''_j(\bar{x}_j) > 0.$$

Now it follows readily from (25) and (27) that $g''(x_0)$ is continuous and positive except possibly at $y_1, \dots, y_{n-1}$.

Next we observe that the only term in the denominator of

$\lim_{x_0 \rightarrow y_h^+} g''(x_0)$ which is not also in the denominator of

$\lim_{x_0 \rightarrow y_h^-} g''(x_0)$ is

$$(28) \quad \lim_{x_0 \rightarrow y_h^+} 1/p_h q''_h(\bar{x}_h) = \lim_{t \rightarrow 0} 1/p_h q''_h(t).$$

This equation establishes the statements about $g''(x_0)$ at $y_1, \dots, y_{n-1}$. Finally, if $0 < x_0 \leq y_{n-1}$ we have

$$(29) \quad g''(x_0) = p_n q''_n(x_0)$$

and the final statement follows from this formula. We have also established the following corollary:

Corollary 1. If $q_1(t), \dots, q_n(t)$ are functions of type A and if $\lim_{t \rightarrow 0^+} q_i''(t) = \infty$ ($i=1, \dots, n$) then $g(x_0)$ is also of type A with $\lim_{x_0 \rightarrow 0^+} g''(x_0) = \infty$.

§ 5. Functions of Type B.

If conditions A are replaced by the weaker conditions

- (B) $\left\{ \begin{array}{l} \text{(i), (iii), (iv) same as for A} \\ \text{(ii): } q''(x_0) > 0 \text{ for all } t > 0 \text{ except for a finite} \\ \text{number of points and at these points the one} \\ \text{sided limits exist and are positive; and } q'(x_0) \\ \text{is continuous for all } t > 0. \end{array} \right.$

we have a new class of functions which we call functions of type B. Theorem 2 states essentially that $g(x_0)$ is a function of type B. Actually, if the initial $q_i(t)$ are all of type B the conclusions of Theorems 1 and 2 still hold except that $g''(x_0)$ will have as points of discontinuity not only the y_j but also those x_0 for which there exists a j such that $\bar{x}_j$ is a point of discontinuity of $q_j''(t)$.

§ 6. The Inverse Power Case.

There are various ways in which the functions $q_i(t)$ of type A may be encountered. If $p(t)$ is a function which satisfies A(i), (ii), (iii) and if s is any positive vector, then the functions

$$(30) \quad q_j(t) = \frac{p(x_j + s_j)}{p(s_j)} \quad (j=1, \dots, n)$$

are of type A.

In general one cannot expect explicit solutions for the y_j and $\mu_j(x_0)$. However in the special case

$$(31) \quad p(t) = t^{-\gamma} \quad (\gamma > 0)$$

of (30) it is possible to obtain explicit solutions. In this case we have

$$(32) \quad \begin{cases} q_j(x_j) = s_j^\gamma / (s_j + x_j)^\gamma \\ q'_j(x_j) = -\gamma s_j^\gamma / (s_j + x_j)^{\gamma+1} \end{cases} \quad (j=1, \dots, n)$$

and hence

$$(33) \quad \begin{cases} \rho_j = -\gamma p_j / s_j \\ r_j(w) = {}^{\gamma+1} \sqrt{-\gamma s_j^\gamma / w} - s_j \end{cases} \quad (j=1, \dots, n)$$

Then solving (15) for μ_h we get

$$(34) \quad \mu_h = \mu_h(x_0) = -(\tau_h / (x_0 + \sigma_h))^{\gamma+1} \quad (h=1, \dots, n)$$

where

$$(35) \quad \begin{cases} \tau_h = \sum_{j=h}^n {}^{\gamma+1} \sqrt{\gamma s_j^\gamma p_j} \\ \sigma_h = \sum_{j=h}^n s_j \end{cases}$$

Next, from (34) with $\mu_h = \rho_h$, $x_0 = y_h$ we get

$$(36) \quad y_h = -\sigma_h + \tau_h^{\gamma+1} \sqrt{-\rho_h} \quad (h=1, \dots, n).$$

If h is now determined according to (18) we get for the maximizing vector $\bar{x}$

$$(37) \quad \begin{cases} \bar{x}_j = 0 & (j < h) \\ \bar{x}_j = \frac{(x_0 + \sigma_h)^{\gamma+1}}{\tau_h} \sqrt{\gamma s_j^\gamma p_j} - s_j & (j \geq h) \end{cases}$$

and

$$(38) \quad g(x_0) = \sum_{j=1}^{h-1} p_j + \frac{\tau_h^{\gamma+1}}{\gamma(x_0 + \sigma_h)^\gamma}.$$

§7. A Maximization Problem.

Let x, p and $f(x, p)$ be as in §4. We consider a function

$$(39) \quad F(x, p) = K(x_0, f(x, p))$$

where for each x_0 $K(x_0, t)$ is monotone decreasing in t . Thus to maximize F for x in $\Gamma(x_0)$ we merely choose the $\bar{x}$ which minimizes f . However, if the maximization is to be for all x with $x \geq 0$ or for all x with $x_0 \leq w$, one can proceed as follows. For each x_0 let $G(x_0)$ be the maximum of F for

$x \in \Gamma(x_0)$; i.e.,

$$(40) \quad G(x_0) = K(x_0, g(x_0)).$$

Then the maximization of $F(x, p)$ is reduced to the scalar problem of maximizing $G(x_0)$. If $K(x_0, t)$ satisfies certain differentiability conditions we can make use of the results of Theorem 2 and apply the methods of elementary calculus to achieve the maximization. Of course, the nature of the function $g(x_0)$ will require that each interval $y_h < x_0 \leq y_{h-1}$, be studied separately.

One choice of $K(x_0, t)$ has been treated in RM-909 and also by R. Isaacs. This is the case of (31) with $\gamma = 1$ and

$$(41) \quad K(x_0, t) = Q(x_0 + s_0)(1-t) - x_0,$$

where Q is a constant in the interval $0 < Q \leq 1$. Interpretations for this case were given in RM-909.

A more general case is

$$(42) \quad K(x_0, t) = H(x_0)(1-t) - x_0$$

where $H(z)$ satisfies the following conditions:

$$(43) \quad \left\{ \begin{array}{ll} \text{(i)} & H(z) \geq 0 \quad 0 \leq z < \infty \\ \text{(ii)} & H(z) - z \text{ is bounded from above for } z \geq 0 \\ \text{(iii)} & H(z) \text{ and } H'(z) \text{ are continuous for } z \geq 0. \end{array} \right.$$

These conditions are sufficient to guarantee the existence of a maximum of $G(x_0) = K(x_0, g(x_0))$. Since $g(0) = 1$, $G(0) = 0$. A sufficient condition for a positive maximum is then $G'(0) > 0$, and this maximum will occur at a point $\bar{x}_0$ for which $G'(\bar{x}_0) = 0$. If K is given by (42), we have

$$(44) \quad \begin{cases} G(x_0) = H(x_0)(1-g(x_0)) - x_0 \\ G'(x_0) = H'(x_0)(1-g(x_0)) - H(x_0)g'(x_0) - 1 \\ G''(x_0) = H''(x_0)(1-g(x_0)) - 2H'(x_0)g'(x_0) - H(x_0)g''(x_0), \end{cases}$$

where the formula for $G''(x_0)$ holds only when $H''(x_0)$ and $g''(x_0)$ exist.

The case

$$(45) \quad H(t) = c = \text{constant}$$

is easily handled. Then $G'(0) = -c\rho_n - 1$. If $c \leq -1/\rho_n$ then $x_0 = 0$ is the only maximum. If $c > -1/\rho_n$, we determine h so that

$$(46) \quad \rho_h < -\frac{1}{c} \leq \rho_{h-1}$$

and then select $\bar{x}_0$ so that $\mu_h(\bar{x}_0) = -\frac{1}{c}$. This will be the maximizing value. This case is the only one where I have been able to find an explicit solution for arbitrary functions of the case $q_j(t)$.

$$(47) \quad H(x_0) = Q \cdot (x_0 + s_0), \quad q_j(t) = s_j^\gamma / (s_j + x_j)^\gamma$$

has been mentioned above for $\gamma = 1$. We now consider the case for $\gamma > 1$.

We have

$$(48) \quad G''(x_0) = \frac{Q \tau_h^{\gamma+1}}{(x_0 + \sigma_h)^{\gamma+1}} \left[2 - \frac{(\gamma+1)(x_0 + s_0)}{(x_0 + \sigma_h)} \right].$$

Since $\gamma > 1$ and $s_0 \geq \sigma_h$ we have $G''(x_0) < 0$ for $x_0 > 0$.

Hence there is a unique maximum. A simple calculation shows that

$$(49) \quad G'(0) = \gamma Q s_0 p_n / s_n - 1$$

and hence that the maximum value of G is positive if and only if

$$(50) \quad Q > s_n / (\gamma s_0 p_n).$$

In general for $y_h \leq x \leq y_{h-1}$ we find that

$$(51) \quad G'(x) = -1 + Q \left\{ \pi_h + \left(1 - \frac{1}{\gamma}\right) \tau_h (-\mu_h)^{\gamma/\gamma+1} + (s_0 - \sigma_h) (-\mu_h) \right\}.$$

If $G'(y_h) \geq 0$ but $G'(y_j) < 0$ for $j < h$, then the maximizing value is in the interval $y_h \leq x < y_{h-1}$ and is given by

$$(52) \quad x_0 = -\sigma_h + \tau_h/t$$

where t is the positive root of

$$(53) \quad t^{\gamma+1} + \frac{\tau_h}{s_0 - \sigma_h} \left(1 - \frac{1}{\gamma}\right) t^\gamma - \frac{1 - Q\pi_h}{Q(s_0 - \sigma_h)} = 0.$$

15/ STI-ATI-210 473

UNCLASSIFIED

18/7 Rand Corp., Santa Monica, Calif.

DECISION PROCESSES AND FUNCTIONAL EQUATIONS, by S.
Karlin and Harold N. Shapiro. 3 Sep 52, 10p. (Rept. no.
RM-933) (Contract AF 33(038)6413)

SUBJECT HEADINGS

DIV: Mathematics (15)

Programming

SECT: All

Strategy

DIV: Military Sciences &

Topology

Operations (18)

SECT: Strategy & Tactics (7)

DIST: Copies obtainable from ASTIA-DSC

Proj. Rand



UNCLASSIFIED