

Lectures on  
OPTIMIZATION — THEORY AND ALGORITHMS

By  
JEAN CEA



TATA INSTITUTE OF FUNDAMENTAL RESEARCH  
BOMBAY

1978

0224  
C2

8063456

Lectures on  
**OPTIMIZATION — THEORY AND ALGORITHMS**

By  
**JEAN CEA**



E8063456

Notes by  
**M. K. V. MURTHY**



Published for the

**TATA INSTITUTE OF FUNDAMENTAL RESEARCH, BOMBAY**



**SPRINGER-VERLAG**

**Berlin Heidelberg New York**

1978

8083458

LOCATED BY

OPTIMIZATION - THEORY AND ALGORITHMS

© TATA INSTITUTE OF FUNDAMENTAL RESEARCH, 1978

---

ISBN 3-540-08850-4 Springer-Verlag Berlin .Heidelberg .New York

ISBN 0-387-08850-4 Springer-Verlag New York .Heidelberg .Berlin

---

No part of this book may be reproduced in any form  
by print, microfilm or any other means without written  
permission from the Tata Institute of Fundamental  
Research, Colaba, Bombay 400 005

Printed by K. P. Puthran at the Tata Press Limited,  
414 Veer Savarkar Marg, Bombay 400 025 and published by H. Goetze,  
Springer-Verlag, Heidelberg, West Germany

PRINTED IN INDIA

TATA INSTITUTE OF FUNDAMENTAL RESEARCH  
LECTURES ON MATHEMATICS AND PHYSICS

General Editors: K.G. RAMANATHAN and B.V. SREEKANTAN

MATHEMATICS

1. On the Riemann Zeta-function. By K. Chandrasekharan\*
2. On Analytic Number Theory. By H. Rademacher\*
3. On Siegel's Modular Functions. By H. Maass\*
4. On Complex Analytic Manifolds. By L. Schwartz\*
5. On the Algebraic Theory of Fields. By K.G. Ramanathan\*
6. On Sheaf Theory. By C.H. Dowker\*
7. On Quadratic Forms. By C.L. Siegel\*
8. On Semi-group Theory and its application to Cauchy's problem in Partial Differential Equations. By K. Yosida\*
9. On Modular Correspondences. By M. Eichler\*
10. On Elliptic Partial Differential Equations. By J.L. Lions\*
11. On Mixed Problems in Partial Differential Equations and Representations of Semi-groups. By L. Schwartz\*
12. On Measure Theory and Probability. By H.R. Pitt\*
13. On the Theory of Functions of Several Complex Variables. By B. Malgrange\*
14. On Lie groups and Representations of Locally Compact Groups. By F. Bruhat\*
15. On Mean Periodic Functions. By J.P. Kahane\*
16. On Approximation by Polynomials. By J.C. Burkill\*
17. On Meromorphic Functions. By W.K. Hayman\*
18. On the Theory of Algebraic Functions of One Variable. By M. Deuring\*

19. On Potential Theory. By M. Brelot\*
20. On Fibre Bundles and Differential Geometry. By J. L. Koszul\*
21. On Topics in the Theory of Infinite Groups. By B. H. Neumann\*
22. On Topics in Mean Periodic Functions and the Two-Radius Theorem.  
By J. Delsarte\*
23. On Advanced Analytic Number Theory. By C. L. Siegel\*
24. On Stochastic Processes. By K. Ito
25. On Exterior Differential Systems. By M. Kuranishi\*
26. On Some Fixed Point Theorems of Functional Analysis. By F. F. Bonsall\*
27. On Some Aspects of p-adic Analysis. By F. Bruhat\*
28. On Riemann Matrices. By C. L. Siegel\*
29. On Modular Functions of one complex variable. By H. Maass
30. On Unique Factorization Domains. By P. Samuel\*
31. On The Fourteenth Problem of Hilbert. By M. Nagata
32. On Groups of Transformations. By J. L. Koszul
33. On Geodesics in Riemannian Geometry. By M. Berger\*
34. On Topics in Analysis. By Raghavan Narasimhan\*
35. On Cauchy Problem. By Sigeru Mizohata
36. On Old and New Results on Algebraic Curves. By P. Samuel\*
37. On Minimal Models and Birational Transformations of Two Dimensional  
Schemes. By I. R. Shafarevich
38. On Stratification of Complex Analytic Sets. By M. -H. Schwartz
39. On Levi Convexity of Complex Manifolds and Cohomology Vanishing  
Theorems. By E. Vesentini
40. On An Introduction to Grothendieck's Theory of the Fundamental  
Group. By J. P. Murre

41. On Topics in Algebraic K-Theory. By Hyman Bass\*
42. On The Singularities of the three-body problem. By C.L. Siegel
43. On Polyhedral Topology. By J.R. Stallings
44. On Introduction to Algebraic Topology. By G. de Rham
45. On Quadratic Jordan Algebras. By N. Jacobson
46. On The Theorem of Browder and Novikov and Siebenmann's Thesis.  
By M.A. Kervaire
47. On Galois Cohomology of Classical Groups. By M. Kneser
48. On Discrete subgroups of Lie groups. By G.D. Mostow
49. On The Finite element method. By Ph. Ciarlet
50. On Disintegration of measures. By L. Schwartz
51. On Introduction to moduli problems and orbit spaces. By P.E. Newstead
52. On Numerical methods for time dependent equations - Application to  
Fluid flow problems. By P. Lascaux
53. On Optimisation - Theory and Algorithms. By Jean Cea
54. On Deformations of Singularities. By M. Artin
55. On Sieve Methods. By Hans E. Richert
56. On Irregularities of Distribution. By Wolfgang M. Schmidt
57. On Expansion techniques in Algebraic Geometry. By S.S. Abhyankar
58. On Torus Embeddings and Applications. By T. Oda

# CONTENTS

## CHAPTER I - DIFFERENTIAL CALCULUS IN NORMED LINEAR SPACES

1. Gateaux derivatives	...	1
2. Taylor's formula	...	7
3. Convexity and Gateaux differentiability	...	11
4. Gateaux differentiability and weak lower semi-continuity		14
5. Commutation of derivations	...	
6. Frechet derivatives	...	6
7. Model problem	...	19

## CHAPTER II - MINIMIZATION OF FUNCTIONALS-THEORY

1. Minimization without convexity conditions	...	21
2. Minimization with convexity conditions	...	25
3. Applications to the model problem and reduction to variational inequality	...	30
4. Some functional spaces	...	33
5. Examples	...	36

## CHAPTER III - MINIMIZATION WITHOUT CONSTRAINTS - ALGORITHMS

1. Method of descent : (1.1) Generalities	...	51
	(1.2) Convergent choices of the direction of descent $w_k$ ...	54
	(1.3) Convergent choices of $\rho_k$	59
	(1.4) Convergence of algorithms	69
2. Generalized Newton's method	...	73
3. Other methods	...	87



## CHAPTER IV - MINIMIZATION WITH CONSTRAINTS- ALGORITHMS

	...	88
1. Linearization method	...	88
2. Centre method	...	105
3. Method of gradient with projection	...	109
4. Minimization in a product space	...	113
(4.1) Statement of the problem	...	113
(4.2) Minimization with constraints of convex functionals on products of reflexive Banach spaces	...	116
(4.3) Main result-convergence of the algorithm and Gauss-Seidel method	...	118
(4.4) Some applications-differentiable and non- differentiable functionals in finite dimensions		124
(4.5) Minimization of quadratic functionals on Hilbert spaces - Relaxation method by blocks	...	126
(4.6) Algorithm (of relaxation method)-Details	...	128
(4.7) Convergence of the algorithm	...	130
(4.8) Some examples of relaxation method in finite dimensional spaces	...	139
(4.9) Examples in infinite dimensional Hilbert spaces- optimization with constraints in Sobolev spaces		142
CHAPTER V - <u>DUALITY AND ITS APPLICATIONS</u>	...	144
1. Preliminaries - Recollection of Hahn-Banach and Ky Fan and Sion theorem, Lagrangian and Lagrange multipliers, Primal and dual problems	...	145
2. Duality in finite dimensional spaces via Hahn-Banach theorem	...	155
Qualifying hypothesis	...	159
Some examples of qualifying hypothesis	...	160



3.	Duality in infinite dimensional spaces via Ky Fan and Sion theorem	...	162
(3.1)	Duality in the case of a quadratic form	...	163
(3.2)	Dual problem	...	170
(3.3)	Method of Uzawa	...	173
4.	Minimization of non-differentiable functionals using duality - examples and algorithm	...	183
VI -	<u>ELEMENTS OF THEORY OF OPTIMAL CONTROL AND ELEMENTS OF OPTIMAL DESIGN</u>	...	194
1.	Optimal control theory	...	194
(1.1)	Formulation of the problem of optimal control		196
(1.2)	Duality and existence	...	198
(1.3)	Elimination of state	...	207
(1.4)	Approximation	...	211
2.	Theory of optimal design	...	215
(2.1)	Formulation of the problem of optimal design		217
(2.2)	A simple example	...	221
(2.3)	Computation of the derivative of the cost function		222
(2.4)	Hypothesis and results	...	226
	BIBLIOGRAPHY	...	233

## CHAPTER I

### DIFFERENTIAL CALCULUS IN NORMED LINEAR SPACES

We shall recall in this chapter the notions of differentiability in the sense of Gateaux and Frechet for mappings between normed linear spaces and some of the properties of derivatives in relation to convexity and weak lower semi-continuity of functionals on normed linear spaces. We shall use these concepts throughout our discussions.

In the following all the vector spaces considered will be over the field of real numbers  $\mathbb{R}$ .

If  $V$  is a normed (vector) space we shall denote by  $\|\cdot\|_V$  the norm in  $V$ , by  $V'$  its (strong) dual with  $\|\cdot\|_{V'}$  as the norm and by  $\langle \cdot, \cdot \rangle_{V' \times V}$  the duality pairing between  $V$  and  $V'$ . If  $V$  is a Hilbert space then  $(\cdot, \cdot)_V$  will denote the inner product in  $V$ . If  $V$  and  $H$  are two normed spaces then  $\mathcal{L}(V, H)$  denotes the vector space of all continuous linear mappings from  $V$  into  $H$  provided with the norm  $A \rightarrow \|A\|_{\mathcal{L}(V, H)} = \sup \{ \|Av\|_H / \|v\|_V, v \in V \}$ .

#### 1. GATEAUX DERIVATIVES

Let  $V, H$  be normed spaces and  $A : U \subset V \rightarrow H$  be a mapping of an open subset  $U$  of  $V$  into  $H$ . We shall often call a vector  $\varphi \in V, \varphi \neq 0$  a direction in  $V$ .

DEFINITION 1.1. The mapping  $A$  is said to be differentiable in the sense of Gateaux or simply  $G$ -differentiable at a point  $u \in U$  in the direction  $\varphi$  if the difference quotient

$$(A(u + \theta\varphi) - A(u))/\theta$$

has a limit  $A'(u, \varphi)$  in  $H$  as  $\theta \rightarrow 0$  in  $\mathbb{R}$ . The (unique) limit  $A'(u, \varphi)$  is called the Gateaux derivative of  $A$  at  $u$  in the direction  $\varphi$ .

$A$  is said to be  $G$ -differentiable in a direction  $\varphi$  in a subset of  $U$  if it is  $G$ -differentiable at every point of the subset in the direction  $\varphi$ .

We shall simply call  $A'(u, \varphi)$  the  $G$ -derivative of  $A$  at  $u$  since the dependence on  $\varphi$  is clear from the notation.

REMARK 1.1. The operator  $V \ni \varphi \mapsto A'(u, \varphi) \in H$  is homogeneous:

$$A'(u, \alpha \varphi) = \alpha A'(u, \varphi) \text{ for } \alpha > 0.$$

In fact,

$$A'(u, \alpha \varphi) = \lim_{\theta \rightarrow 0} (A(u + \alpha \theta \varphi) - A(u)) / \theta = \alpha \lim_{\lambda \rightarrow 0} (A(u + \lambda \varphi) - A(u)) / \lambda = \alpha A'(u, \varphi).$$

However, this operator is not, in general, linear as can be seen immediately from Example 1.2 below.

We shall often denote a functional on  $U$  by  $J$ .

REMARK 1.2. Every linear functional  $L: V \rightarrow \mathbb{R}$  is  $G$ -differentiable everywhere in  $V$  in all directions and its  $G$ -derivative is

$$L'(u, \varphi) = L(\varphi)$$

since  $(L(u + \theta \varphi) - L(u)) / \theta = L(\varphi)$ . It is a constant functional (i. e. independent of  $u$  in  $V$ ).

If  $a(u, v): V \times V \rightarrow \mathbb{R}$  is a bilinear functional on  $V$  then the functional  $J: V \ni v \mapsto J(v) = a(v, v) \in \mathbb{R}$  is  $G$ -differentiable everywhere in all directions and

$$J'(u, \varphi) = a(u, \varphi) + a(\varphi, u).$$

If further  $a(u, v)$  is symmetric (i. e.  $a(u, v) = a(v, u)$  for all  $u, v \in V$ ) then

$J'(u, \varphi) = 2a(u, \varphi)$ . This follows immediately from bilinearity:

$$a(u + \theta\varphi, u + \theta\varphi) = a(u, u) + \theta(a(u, \varphi) + a(\varphi, u)) + \theta^2 a(\varphi, \varphi)$$

so that

$$J'(u, \varphi) = \lim_{\theta \rightarrow 0} (J(u + \theta\varphi) - J(u)) / \theta = a(u, \varphi) + a(\varphi, u).$$

The following example will be a model case of linear problems in many of our discussions in the following chapters.

**Example 1.1.** Let  $(u, v) \mapsto a(u, v)$  be a symmetric bi-linear form on a Hilbert space  $V$  and  $v \mapsto L(v)$  a linear form on  $V$ . Define the functional  $J : V \rightarrow \mathbb{R}$  by

$$J(v) = \frac{1}{2} a(v, v) - L(v).$$

It follows from the above Remark that  $J$  is  $G$ -differentiable everywhere in  $V$  in all directions  $\varphi$  and

$$J'(u, \varphi) = a(u, \varphi) - L(\varphi).$$

In many of the questions we shall assume:

(i)  $a(\cdot, \cdot)$  is (bi-) continuous: there exists a constant  $M > 0$  such that

$$a(u, v) \leq M \|u\|_V \|v\|_V \quad \text{for all } u, v \in V;$$

(ii)  $a(\cdot, \cdot)$  is  $V$ -coercive: there exists a constant  $\alpha > 0$  such that

$$a(v, v) \geq \alpha \|v\|_V^2 \quad \text{for all } v \in V$$

and

(iii)  $L$  is continuous: there exists a constant  $N > 0$  such that

$$L(v) \leq N \|v\|_V \quad \text{for all } v \in V.$$

Example 1.2. The function  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \begin{cases} 0 & \text{if } (x, y) = (0, 0) \\ x^5 / ((x-y)^2 + x^4) & \text{if } (x, y) \neq (0, 0) \end{cases}$$

is  $G$ -differentiable everywhere and in all directions. In fact, if  $u = (0, 0) \in \mathbb{R}^2$  then given a direction  $\varphi = (X, Y) \in \mathbb{R}^2$  ( $\varphi \neq 0$ ) we have

$$(f(\theta X, \theta Y) - f(0, 0)) / \theta = \theta^2 X^5 / ((X-Y)^2 + \theta^2 X^4)$$

which has a limit as  $\theta \rightarrow 0$  and we have

$$f'(u, \varphi) = f'((0, 0), (X, Y)) = \begin{cases} 0 & \text{if } X \neq Y \\ X & \text{if } X = Y \end{cases}$$

One can also check easily that  $f$  is  $G$ -differentiable in  $\mathbb{R}^2$ .

The following will be the general abstract form of functionals in many of the non-linear problems that we shall consider.

Example 1.3. Let  $\Omega$  be an open set in  $\mathbb{R}^n$  and  $V = L^p(\Omega)$ ,  $p > 1$ . Suppose  $g: \mathbb{R}^1 \ni t \mapsto g(t) \in \mathbb{R}^1$  be a  $C^1$ -function such that

$$(i) \quad |g(t)| \leq C|t|^p \quad \text{and} \quad (ii) \quad |g'(t)| \leq C|t|^{p-1}$$

for some constant  $C > 0$ . Then

$$u \mapsto J(u) = \int_{\Omega} g(u(x)) \, dx$$

defines a functional  $J$  on  $L^p(\Omega) = V$  which is  $G$ -differentiable everywhere in all directions and we have

$$J'(u, \varphi) = \int_{\Omega} g'(u(x)) \varphi(x) dx.$$

(The right hand side here exists for any  $u, \varphi \in L^p(\Omega)$ ).

In fact, since  $u \in L^p(\Omega)$  and since  $g$  satisfies (i) we have

$$|J(u)| \leq \int_{\Omega} |g(u)| dx \leq C \int_{\Omega} |u|^p dx < +\infty$$

which means  $J$  is well defined on  $L^p(\Omega)$ . On the other hand, for any  $u \in L^p(\Omega)$ , since  $g'$  satisfies (ii),  $g'(u) \in L^{p'}(\Omega)$  where  $p^{-1} + p'^{-1} = 1$ . For, we have

$$\int_{\Omega} |g'(u)|^{p'} dx \leq C \int_{\Omega} |u|^{(p-1)p'} dx = C \int_{\Omega} |u|^p dx < +\infty.$$

Hence, for any  $u, \varphi \in L^p(\Omega)$ , we have by Hölder's inequality

$$\left| \int_{\Omega} g'(u) \varphi dx \right| \leq \|g'(u)\|_{L^{p'}(\Omega)} \|\varphi\|_{L^p(\Omega)} \leq C \|u\|_{L^p(\Omega)}^{p/p'} \|\varphi\|_{L^p(\Omega)} < +\infty.$$

To compute  $J'(u, \varphi)$ , if  $\theta \in \mathbb{R}$  we define  $h : [0, 1] \rightarrow \mathbb{R}$  by setting

$$h(t) = g(u + t\theta\varphi).$$

Then  $h \in C^1(0, 1)$  and

$$h(1) - h(0) = \int_0^1 h'(t) dt = \theta \varphi(x) \int_0^1 g'(u + t\theta\varphi) dt$$

( $t = t(x), |t(x)| \leq 1$ ) so that

$$(J(u + \theta\varphi) - J(u)) / \theta = \int_{\Omega} \varphi(x) \int_0^1 g'(u(x) + t\theta\varphi(x)) dt dx.$$

One can easily check as above that the function

$$(x, t) \longmapsto \varphi(x) g'(u(x) + t\theta\varphi(x))$$

belongs to  $L^1(\Omega \times [0, 1])$  and hence by Fubini's theorem

$$(J(u+\theta\varphi) - J(u))/\theta = \int_0^1 dt \int_{\Omega} \varphi(x) g'(u(x) + t\theta\varphi(x)) dx.$$

Here the continuity of  $g'$  implies that

$$g'(u+t\theta\varphi) \longrightarrow g'(u) \text{ as } \theta \longrightarrow 0 \text{ (and hence as } t\theta \longrightarrow 0)$$

uniformly for  $t \in [0, 1]$ . Moreover, the condition (ii) together with triangle inequality implies that, for  $0 < \theta \leq 1$ ,

$$|\varphi(x) g'(u(x) + t\theta\varphi(x))| \leq C |\varphi(x)| (|u(x)| + |\varphi(x)|)^{p-1}$$

and the right side is integrable by Hölder's inequality. Then by dominated convergence theorem we conclude that

$$J'(u, \varphi) = \int_{\Omega} g'(u) \varphi dx.$$

DEFINITION 1.2. An operator  $A : U \subset V \longrightarrow H$  ( $U$  being an open set in  $V$ ) is said to be twice differentiable in the sense of Gateaux at a point  $u \in V$  in the directions  $\varphi, \Psi$  ( $\varphi, \Psi \in V, \varphi \neq 0, \Psi \neq 0$  given) if the operator  $u \mapsto A'(u, \varphi) : U \subset V \longrightarrow H$  is once G-differentiable at  $u$  in the direction  $\Psi$ . The G-derivative of  $u \mapsto A'(u, \varphi)$  is called the second G-derivative of  $A$  and is denoted by  $A''(u, \varphi, \Psi) \in H$ .

$$\text{i.e.} \quad A''(u; \varphi, \Psi) = \lim_{\theta \rightarrow 0} (A'(u + \theta\Psi, \varphi) - A'(u, \varphi))/\theta.$$

REMARK 1.3. Derivatives of higher orders in the sense of Gateaux can be defined in the same way. As we shall not use derivatives of higher orders in the following we shall not consider their properties.

Now let  $J : U \subset V \longrightarrow \mathbb{R}$  be a functional on an open set of a normed linear space  $V$  which is once G-differentiable at a point  $u \in U$ . If the functional



$\varphi \mapsto J'(u, \varphi)$  is continuous linear on  $V$  then there exists a (unique) element  $G(u) \in V'$  such that

$$J'(u, \varphi) = \langle G(u), \varphi \rangle_{V' \times V} \text{ for all } \varphi \in V.$$

Similarly, if  $J$  is twice  $G$ -differentiable at a point  $u \in U$  and if the form  $(\varphi, \psi) \mapsto J''(u; \varphi, \psi)$  is a bilinear (bi-) continuous form on  $V \times V$  then there exists a (unique) element  $H(u) \in \mathcal{L}(V, V')$  such that

$$J''(u; \varphi, \psi) = \langle H(u)\varphi, \psi \rangle_{V' \times V}.$$

DEFINITION 1.3.  $G(u) \in V'$  is called the gradient of  $J$  at  $u$  and  $H(u) \in \mathcal{L}(V, V')$  is called the Hessian of  $J$  at  $u$ .

## 2. TAYLOR'S FORMULA

We shall next deduce the mean value theorem and Taylor's formula of second order for a mapping  $A : U \subset V \longrightarrow H$  ( $U$  open subset of a normed linear space  $V$ ) in terms of the  $G$ -derivatives of  $A$ . We shall begin with the case of functionals on a normed linear space  $V$ .

Let  $J$  be a functional defined on an open set  $U$  in a normed linear space  $V$  and  $u, \varphi \in V, \varphi \neq 0$  be given. Throughout this section we assume that the set  $\{u + \theta\varphi; \theta \in [0, 1]\}$  is contained in  $U$ . It is convenient to introduce the function  $f : [0, 1] \longrightarrow \mathbb{R}$  by setting

$$\theta \longrightarrow f(\theta) = J(u + \theta\varphi).$$

We observe that if  $J'(u + \theta\varphi, \varphi)$  exists then  $f$  is once differentiable in  $]0, 1[$  and, as one can check immediately

$$f'(\theta) = J'(u + \theta \varphi, \varphi).$$

Similarly if  $J''(u + \theta \varphi, \varphi, \varphi)$  exists then  $f$  is twice differentiable and

$$f''(\theta) = J''(u + \theta \varphi; \varphi, \varphi).$$

**PROPOSITION 2.1.** Let  $J$  be a functional on an open set  $U$  of a normed space  $V$  and  $u \in U, \varphi \in V$  be given. If  $\{u + \theta \varphi; \theta \in [0, 1]\} \subset U$  and  $J$  is once  $G$ -differentiable on this set in the direction  $\varphi$  then there exists a  $\theta_0 \in ]0, 1[$  such that

$$(2.1) \quad J(u + \varphi) = J(u) + J'(u + \theta_0 \varphi, \varphi)$$

**Proof.** This follows immediately from the classical mean value theorem applied to the function  $f$  on  $[0, 1]$ : there exists a  $\theta_0 \in ]0, 1[$  such that

$$f(1) = f(0) + 1 \cdot f'(\theta_0)$$

which is nothing but (2.1).

**PROPOSITION 2.2.** Let  $U$  and  $J$  be as in Proposition 2.1. If  $J$  is twice  $G$ -differentiable on the set  $\{u + \theta \varphi; \theta \in [0, 1]\}$  in the directions  $\varphi, \varphi$  then there exists a  $\theta_0 \in ]0, 1[$  such that

$$(2.2) \quad J(u + \varphi) = J(u) + J'(u, \varphi) + \frac{1}{2} J''(u + \theta_0 \varphi; \varphi, \varphi).$$

This again follows from the classical Taylor's formula applied to the function  $f$  on  $[0, 1]$ .

**REMARK 2.1.** If  $L: V \rightarrow \mathbb{R}$  is a linear functional on  $V$  then by Remark 1.1 is  $G$ -differentiable everywhere in all directions and we find that the formula (2.1) reads

$$L(u + \varphi) = L(u) + L(\varphi)$$