Exercises for Optimization course

Manuel Samuelides
manuel.samuelides@isae.fr

10th of October 2010

Contents

1 Unconstrained Optimization:Gradient methods (C01-C02) 1
1.1 Optimization of a compact level set function 1
1.2 Taylor formula and minimization 2
1.3 Convexity and differentiability 2
1.4 Practical analysis 2
1.5 Orthogonality of steepest descent directions. 2
1.6 Gradient and conjugate gradient optimizations 3

2 Linear programming (C04) 3

3 Constrained optimization (C10) 3
3.1 Elementary quadratic programming 3
3.2 Lagrange multipliers as sensitivities 3
3.3 Kuhn-Tucker theorem practical application 4
3.4 Optimal control within finite time 7

1 Unconstrained Optimization:Gradient methods (C01-C02)

1.1 Optimization of a compact level set function

Prove the following theorem using the classical theorem about continuous functions on compact set:

Theorem 1 *Let J a real continuous compact level set function on Ω. There exists $x^\star \in \Omega$ which is a **minimizer** of J i.e. $\forall x \in \Omega, J(x^*) \leq J(x)$*

1.2 Taylor formula and minimization

Use Taylor approximation formula of first and second order:

- $J(x+h) = J(x) + (\nabla_x J \mid h) + o(\| h \|)$
- $J(x+h) = J(x) + (\nabla_x J \mid h) + \frac{1}{2}(h \mid \nabla_x^2 J\, h) + o(\| h \|^2)$

to prove the following:

1. If $J \in \mathcal{C}^1(\Omega)$, any minimizer $x^\star$ of J is a stationary point i.e. $\nabla_{x^\star} J = 0$.
2. If $J \in \mathcal{C}^2(\Omega)$ and if $x^\star$ is a minimizer of J, then $\nabla_{x^\star}^2 J$ is a non negative symmetric matrix (eigenvalues are positive or null).
3. Build a real compact level set differentiable function J of two variables such that 0 is stationary point which is neither a minimizer of J nor of $-J$. Such points are called **saddle points**.

1.3 Convexity and differentiability

Show the following properties of a $\mathcal{C}^2(\Omega)$ convex function J using Taylor approximation formula:

1. The hessian of J is non-negative at any point of the domain .
2. A stationary point of a J is a minimizer.

1.4 Practical analysis

1. Discuss compact level set property of the function $J(x, y) = 100(y - x^2)^2 + (1 - x)^2$.
2. Show this function has only one minimizer.
3. Check the function is elliptic in the neighbourhood of its minimizer by computing the Hessian.

1.5 Orthogonality of steepest descent directions.

Prove that the successive descent directions of the steepest descent algorithms are orthogonal.

Solution

Suppose that at step $n-1$, the current point is x_{n-1} and we move according to the steepest descent direction $u_n = -\nabla J(x_{n-1})$.

So we perform the line search: $\min_{t>0} J(x_{n-1} + hu_n)$.
The derivative of the target function is $(\nabla J(x_{n-1} + hu_n) \mid u_n)$.
We stop at the stationary minimizer h^* which checks $(\nabla J(x_{n-1}+h^* u_n) \mid u_n) = 0$
Then we set $x_n = x_{n-1} - h^* u_n, \qquad u_{n+1} = -\nabla J(x_n)$.
So we have $(u_n + 1 \mid u_n) = 0$.

1.6 Gradient and conjugate gradient optimizations

Let us consider $\qquad J(x, y) = \frac{x^2+2y^2}{2}$

1. Let $(x_0; y_0) \in \mathbb{R}^2$. Compute $u_1 = -\nabla_{(x_0,y_0)} J$, h_1 the associate step of the line search in the direction u_1 and (x_1, y_1) the next step that is defined by $(x_1, y_1) = (x_0, y_0) + h_1 u_1$,
2. Then compute $u_2 = -\nabla_{(x_1,y_1)}$, h_2 and x_2.
3. Compute v_2 the conjugate direction of u_1 with respect to the hessian $\nabla^2 J$. Check that the associate linear search allows to reach immediately the minimum of J.

2 Linear programming (C04)

3 Constrained optimization (C10)

3.1 Elementary quadratic programming

Use Lagrange multipliers to project orthogonally 0 in $\mathcal{R}^3$ onto the straight line defined by the following equations

$$\begin{cases} x + 2y - z &= 4 \\ x - y + z &= -\frac{5}{2} \end{cases}$$

3.2 Lagrange multipliers as sensitivities

Let ξ^* be a local minimizer of

$$\min J(\xi)$$

under constraints $c_1(\xi) = 0$ et $c_2(\xi) = 0$. Functions J, c_1, c_2 are $\mathcal{C}^2$ and constraints c_1, c_2 are qualified in ξ^* and locally convex in the neighbourhood of ξ^*.

1. Let us compute the variation of objective for a slight variation of a constraint: $c_1(\xi) = \epsilon$. The solution of the constrained optimzation problem is $\xi(\epsilon)$. Let

us assume that this solution function is derivable with respect to ϵ.
Show that the derivative of the optimal objective value $J[\xi(\epsilon)]$with respect to the constraint ϵ for $\epsilon = 0$ is $-\lambda_1^*$ where λ_1 is the Lagrange parameter associate to the constraint c_1.

2. **Application:** A storage hall with dimensions x, y, z has to be built by co-operative to store 1500 m^3. The building rules oblige to respect the sizing constraint $y = 2z$ (z is the building height, uniform profile of these halls are planned throughout the province).
The buliding cost is 4 C.U. by sq.m. fro walls, 6 for roof and 12 for the ground acquisition. Compute the dimensions of the hall with the needed storage capacity and the minimal cost.
By a careful management, the hall users reduce the storage need for 10 per cent of the capacity. Compute the new dimensions by two methods (straight-forward, using the Lagrange parameter of the first case as a sensitivity parameter and compare the results.

3.3 Kuhn-Tucker theorem practical application

Consider the following constrained minimization:

$$\begin{cases} \max & & x + 4y \\ x & \leq & 4 - y^2 \\ y & \geq & 2x^2 + 1 \end{cases}$$

1. Using Lagrange-Kuhn-Tucker theorem show that there exists at least one solution of the optimization problem.

2. Use a graphical solution to show the solution is unique.

3. Write Kuhn-Tucker equation.

4. Show the first constraint is saturated.

5. Determine explicitly the only solution of the problem

3.4 Optimal control within finite time

Consider the unstationary linear dynamical system

$$\forall t < 0, x_{t+1} = A_{-t}x_t + B_{-t}u_{-t}$$

where (A_t) is a sequence of invertible $(d - d)$matrixes and (B_t) a sequence of vectors in $\mathcal{R}^d$. The end of the evolution is for $t = 0$. We suppose that for any initial state there exists at least a sequence of d scalar controls $(u_{-d}, ..., u_{-1})$ that allows to reach O. The problem for $p > d$ is to find a control trajectory $(u_{-p}, ..., u_{-1})$ which goes from the initial state $x_{-p} = x$ to the final state $x_0 = 0$ minimizing the quadratic cost $J(p) = \frac{1}{2} \sum_{\tau=1}^{p} u_\tau^2$

Notations and hints:

- The unknown variables are the state and control trajectories $[x_{-p}, u_p, ..., x_{-1}, u_1]$, ($2p$ variables)
- The constraints are the initial state,the state evolution equations and the final state ($p + 1$ equations).
- The associate vector Lagrange multipliers are $\lambda_p, ...\lambda_0$

1. Show that the existence and unicity of a solution are straightforward consequences of the equality constrained optimization theorem.
2. Explicit the Lagrangian $\mathcal{L}$ and write the linear system which gives the solution. Compute in function of λ_0 the command trajectory.
3. Show that the application Φ defined by $x \in \mathbb{R}^d \to \Phi(x) = \lambda_0$ is linear and invertible.
4. We define the linear mapping K_t by $x_{-t} = K_t \lambda_0$. K_t does not depend on further solution for time $p > t$. So the problem can be solved by time back-propagation. Determine K_t recursively.
5. *Application:* We consider the kinematics target problem in 2d: $d = 2, A_t = [1, 1; 0, 1], B_t = [\frac{1}{2}; 1]$.
   Compute K_2 and K_3