4.6. AUGMENTED LAGRANGIAN METHODS

Assume all constraints are equalities or the active set of constraints is known.

Problem (NEP): minimize $f(\mathbf{x}) = f(x_1,...,x_n)$
subject to $c_i(\mathbf{x}) = 0, \quad i = 1,...,m$

Let us assume that all problem functions are twice continuously differentiable.

Lagrangian function:

$$L(\mathbf{x}, \mathbf{u}) = f(\mathbf{x}) - \mathbf{u}^T\mathbf{c}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^{m} u_i c_i(\mathbf{x})$$

If $\mathbf{x}^*$ is the solution to the problem (NEP), then $\mathbf{x}^*$ is a stationary point of $L(\mathbf{x}, \mathbf{u}^*)$ but not necessarily the minimum point. The main idea behind the *augmented Lagrangian method* is to modify the Lagrangian so that $\mathbf{x}^*$ becomes a minimizer of this modified function.

Assume differentiation of the Lagrangian (in gradient and Hessian) is always w.r.t. x-variables, Lagrange multipliers are treated as parameters.

<u>Example 4.4.</u> Minimize $f(\mathbf{x}) = x_1^2 - x_2^2$ subject to $x_2 = 0$.

$L(\mathbf{x}, u) = x_1^2 - x_2^2 - ux_2$

$$\nabla L(\mathbf{x},u) = \begin{pmatrix} 2x_1 \\ -2x_2 - u \end{pmatrix}$$

$\nabla L(\mathbf{x}^*,u^*) = 0$ and $x_2 = 0 \Rightarrow \mathbf{x}^* = (0, 0)^T$, $u^* = 0$.

$\mathbf{x}^*$ is a stationary point of $L(\mathbf{x}, u^*)$.

$$\nabla^2 L(\mathbf{x}, u) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$ is indefinite for all $\mathbf{x}$, u $\Rightarrow$ $\mathbf{x}^*$ is a saddle point of $L(\mathbf{x},u^*)$.

Define the *augmented Lagrangian*:

$$L_A(\mathbf{x},\mathbf{u},r) = f(\mathbf{x}) - \sum_{i=1}^{m} u_i c_i(\mathbf{x}) + \frac{1}{2} r \sum_{i=1}^{m} c_i(\mathbf{x})^2$$

(= Lagrangian + quadratic penalty term)

Notation:
$\mathbf{c}(\mathbf{x}) = (c_1(\mathbf{x}), \ldots, c_m(\mathbf{x}))^T$
$\mathbf{a}_i(\mathbf{x}) = \nabla c_i(\mathbf{x})$

$$\mathbf{A}(\mathbf{x}) = \begin{pmatrix} \nabla c_1(\mathbf{x})^T \\ \nabla c_2(\mathbf{x})^T \\ \vdots \\ \nabla c_m(\mathbf{x})^T \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1(\mathbf{x})^T \\ \mathbf{a}_2(\mathbf{x})^T \\ \vdots \\ \mathbf{a}_m(\mathbf{x})^T \end{pmatrix}$$ Jacobian matrix of the constraints

$\mathbf{G_i(x)} = \nabla^2 c_i(\mathbf{x})$ Hessian matrix of $c_i(\mathbf{x})$
$\mathbf{W(x,u)} = \nabla^2 L(\mathbf{x},\mathbf{u})$ Hessian matrix of the Lagrangian

Assume that $\mathbf{A(x^*)}$ has full row rank, i.e. its rows are linearly independent.

$\nabla L_A(\mathbf{x^*},\mathbf{u^*}, r) = \nabla f(\mathbf{x^*}) - \mathbf{A(x^*)}^T\mathbf{u^*} + r\mathbf{A(x^*)}^T\mathbf{c(x^*)} = \nabla L(\mathbf{x^*},\mathbf{u^*}) + r\mathbf{A(x^*)}^T\mathbf{c(x^*)} = \mathbf{0}$,

because the Lagrangian gradient = 0 and $\mathbf{c(x^*)} = \mathbf{0}$.

So $\mathbf{x^*}$ is a stationary point of $L_A(\mathbf{x},\mathbf{u^*},r)$.

$\nabla^2 L_A(\mathbf{x^*},\mathbf{u^*},r) = \mathbf{W(x^*,u^*)} + r\mathbf{A(x^*)}^T\mathbf{A(x^*)}$

$\mathbf{A(x^*)}^T\mathbf{A(x^*)}$ is a positive semidefinite (or positive definite) matrix.

<u>Theorem 4.4.</u>
Assume that the sufficient conditions of Theorem 4.2. hold for $\mathbf{x^*}$, $\mathbf{u^*}$.
There is $r_0 > 0$ such that when $r > r_0$, $\nabla^2 L_A(\mathbf{x^*},\mathbf{u^*},r)$ is positive definite and $\mathbf{x^*}$ is an unconstrained minimum point of $L_A(\mathbf{x},\mathbf{u^*},r)$.

The constrained optimization problem (NEP) can be transformed into an unconstrained optimization problem. Because the optimal Lagrange multipliers are unknown, they have to be estimated iteratively. Like penalty function methods, the augmented Lagrangian method is implemented as sequence of unconstrained minimizations.

Properties of the augmented Lagrangian function
• L_A is differentiable.
• $\mathbf{x^*}$ is the minimizer of $L_A(\mathbf{x},\mathbf{u^*},r)$ with a finite value of r.
• The second order sufficient condition of Theorem 4.2. states that

$$\mathbf{p}^T\mathbf{W(x^*,u^*)p} > 0 \qquad \text{for all } \mathbf{p} \text{ such that } \mathbf{A(x^*)p} = \mathbf{0}.$$

That means the projected Hessian of the Lagrangian, $\mathbf{Z}^T\mathbf{W(x^*,u^*)Z}$ is positive definite.
If this condition doesn't hold, the Hessian of L_A may be indefinite and L_A may be unbounded below.

Augmented Lagrangian method

1. Set intial values for solution point $\mathbf{x_0}$, Lagrange multipliers $\mathbf{u_0}$, penalty parameter r>0, maximum number of iterations kmax. Set k=0.

2. Terminate if $\mathbf{x_k}$ satisfies the optimality conditions or if k > kmax.

3. Minimize $L_A(\mathbf{x},u_k,r)$ with an unconstrained method from starting point $\mathbf{x_k}$. Let $\mathbf{x_{k+1}}$ denote the approximation to the solution.

4. Estimate Lagrange multipliers, using e.g. formula

$$\mathbf{u_{k+1}} = \mathbf{u_k} - r\mathbf{c(x_{k+1})}$$

5. Increase the penalty parameter r, if the constraint violations at $\mathbf{x_{k+1}}$ are not sufficiently smaller than at $\mathbf{x_k}$.

6. Set k=k+1 and go to 2.

Remarks:

• The choice of penalty parameter r is important. When r is too large, the problem becomes ill-conditioned (as in the penalty function method). When r is too small, L_A may be unbounded below or $\mathbf{x}^*$ may not be a local minimum of L_A.

• Estimation of the Lagrange multipliers: the formula comes from the stationarity conditions $\nabla L_A(\mathbf{x}_{k+1},\mathbf{u}_k, r) = \mathbf{0}$ and $\nabla L_A(\mathbf{x}_{k+1},\mathbf{u}_{k+1}, r) = \mathbf{0}$.

The convergence of the method depends on the Lagrange multiplier estimates. The sequence $\mathbf{x}_k$ may converge to the minimum point $\mathbf{x}^*$ only if $\mathbf{u}_k \to u^*$.

• Inequality constraints can be treated with active set methods.

Example 4.4. (continued)

$$\text{Minimize } f(\mathbf{x}) = x_1^2 - x_2^2, \text{ subject to } x_2 = 0$$

using the augmented Lagrangian method. Start with $\mathbf{x}_0 = (1,1)^T$, $u_0 = 1$, $r = 4$.

Example 4.5. Consider the problem of example 4.3:

$$\text{Minimize } f(\mathbf{x}) = x_1 x_2^2, \text{ subject to } 2 - x_1^2 - x_2^2 = 0.$$

The solution is $\mathbf{x}^* \approx (-0.8165, -1.1547)^T$, and the Lagrange multiplier $u^* = 0.8165$.

a) Study the effect of the penalty parameter on the function $L_A(\mathbf{x},u^*,r)$: examine the contour plot with values $r = 0.075$, $r = 0.2$, and $r = 100$.
b) Study the effect of the Lagrange multiplier estimate on the function $L_A(\mathbf{x},u_k,r)$: examine the contour plot with values $u = 0.5$, $u = 0.9$, $u = 1.0$ when the penalty parameter value is set to $r = 0.2$.

The problem and the diagrams can be found in the book Gill, Murray, and Wright: Practical Optimization.