Journal of Environmental Informatics 11(1) 21-27 (2008)

Journal of Environmental Informatics

www.iseis.org/jei

Modified Penalty Function and Parameterization for Solving Least-Cost Treatment of Wastewater

P. Sanongboon* and R. V. Mayorga

Faculty of Engineering, University of Regina, Regina, Saskatchewan, Canada

Received 12 August 2007; revised 24 November 2007; accepted 21 January 2008; published online 31 March 2008

ABSTRACT. This paper proposes a methodology consisting of modified penalty functions and proper parameterization for solving the least-cost treatment of wastewater. Basically, the penalty functions methods are used to handle constraints and can cause very ill-conditioning. With the proper penalty parameters, the ill-conditioned is overcome, a smoother optimization can be obtained, and convergence can be further improved. Once the modified penalty approach is employed, constraints are properly accounted for, and techniques for unconstrained global optimization can be utilized.

Keywords: constrained optimization, global optimum, parameterization, wastewater treatment

1. Introduction

Optimization problems often associate with service, design, planning, scheduling, and so on. A methodology is here presented to deal effectively with both linear and nonlinear constrained optimization problems, and that can provide a global solution. In some situations, the penalty functions are not chosen to deal with constraints because the methods frequently associated with ill-conditioning. Due to ill-conditioning, the transformed unconstrained function it is difficult to optimize. Furthermore, these approaches also tend to exhibit to slow convergence. Instead, constraints are treated as objective functions and techniques for solving multi-objective optimization are utilized (Surry et al., 1995; Fonseca and Fleming, 1995; Loughlin and Ranjithan, 1997). This technique may affect negatively the quality of solutions. Moreover, when global optimum is desired, the problems become even more difficult.

The proposed methodology consists of modifying some penalty functions, and using a parameterization technique for unconstrained optimization for global optimum. By perturbing some parameters over the constraints (Charalambous, 1975, 1976, 1978), the optimum can be easily obtained and the convergence is linearly improved. Under some mild assumptions, the convergence can be further improved by adding a controlling parameter. In particular, by properly setting up the controlling parameter (Mayorga and Quintana, 1982), the convergence can be improved without causing ill-conditioning. Other improvement of the penalty function has been proposed (Mayorga and Quintana, 1985) by adding an auxiliary function to properly scale the approximated Lagrange multipliers for further improvement on the convergence. Also, since there are many parameters involved; updating algorithms for those parameters need to be properly selected.

Since only some methods can guarantee global optimum while others guarantee only local optima, to obtain such global optimum, a method that solves unconstrained optimization via parameterization and inverse function approximation is implemented (Mayorga, 2002; Mayorga and Arriaga, 2002). The method is based on a damped least squares formulation that allows the development of algorithms for unconstrained optimization to attain the global optimum.

2. Modified penalty function

Since equality constraints are always active, the following improvement of penalty functions considers only inequality constraints. Let denote $X = [x_1, x_2, ..., x_n]^T$. The constrained optimization can be written in general form as shown below:

Minimize $F(X)$, (1)

subject to $g_j(X) \geq 0$, for $j = 1, 2, ..., m$. (2)

The concept of the penalty function method is to solve the problems by placing constraints to the objective function. Exterior penalty approach adds penalty parameters for the infeasibility and forces the solution to the feasible region while barrier (or interior) approach adds penalty parameters to ensure that the search will not go out of the feasible region. The

* Corresponding author. Tel.: +1 306 5854726; fax: +1 306 5854855.
E-mail address: sonong1p@uregina.ca (P. Sanongboon).

ISSN: 1726-2135 print/1684-8799 online

following equation is a transformed unconstrained optimization shown below; all constraints are penalized and added into the objective function term.

$$P(X) = F(X) + s\sum_{j=1}^{m} \Phi(g_j(X)) \tag{3}$$

Under some mild assumptions minimizing the above function will give the same solution as the original problem by decreasing close to zero for the interior penalty function and by increasing close to infinity for the exterior penalty function. However, doing this can cause ill-conditioned. Another problem is that it results very slow convergence. To overcome these disadvantages, if the penalty function is rearranged in order not to let the value of tending to zero (or infinity), the very ill-conditioned can be avoided.

A technique for solving equality constrained (Haarhof and Buys, 1970; Hestenes, 1969; Powell, 1969) had been developed to overcome the ill-conditioning and the slow convergence problems. Later the techniques has been extended (Rockafellar, 1974; Fletcher, 1973; Bertsekas, 1974) for the inequalities constrained optimizations.

To overcome the ill-conditioning, an amount $r_j \geq 0$ is perturbed outward the constraints (Charalambous, 1978). Since $\widetilde{X}(s,r)$ is obtained in region $\Re_r$ and $r_i \geq 0$, this gives $X^* \in \Re_r$; therefore $\widetilde{X}(s,r)$ and X^* are contained in the region $\Re_r$. According to this, it is desired that $\widetilde{X}(s,r) \to X^*$. Under mild assumptions the optimum X^* will be a strong minimum of $P(X,s,r)$, which will linearly improve the convergence:

$$P(X,s,r) = F(X) + \sum_{j=1}^{m} s_j \Phi\big(g_j(X) + r_j\big) \tag{4}$$

Other modifications of penalty functions (Mayorga and Quintana, 1982) were introduced for more general modifications of the penalty functions. It resulted that the ill-conditioning was eliminated and convergence was further improved:

$$P(X,s,r) = F(X) + \sum_{j=1}^{m} s_j \Phi\big(h(\cdot) g_j(X) + r_j\big) \tag{5}$$

For the modified exterior penalty function, an amount $r_i \geq 0$ is perturbed inward for each constraint outside region $\Re_t$. The modified exterior penalty function becomes:

$$P(X,s,r) = F(X) + \sum_{j=1}^{m} s_j \Phi\big(h(\cdot) g_j(X) - r_j\big) \tag{6}$$

By choosing the proper parameters, the penalty function represents the original problem and also the approximated Lagrange multipliers are properly scaled. The ill-conditioned is overcome and the further convergence is linearly improved. Note that functions $\Phi(u)$ are $1/u$ and $\log(u)$ are often used for the interior method and functions max $[0, -u]$ and (max $[0, -u])^2$ are often used for the exterior method. For both penalty methods, the function can be selected as:

$$h(\cdot) = \tau_j \text{, for } \tau_j \geq 0 \tag{7}$$

such that:

$$H(\overline{X})(\widetilde{X} - \overline{X}) - \nabla P(\widetilde{X}) = 0 \tag{8}$$

where *H(X)* is a Hessian matrix of *P*(*X, s, r*).

To update the penalty parameters, for the interior method, the following algorithm may be employed:

$$s_j^{k+1} = 0.1 s_j^k \quad \text{as} \quad k \to \infty$$

$$r_j^{k+1} = r_j^k + \tau_j^k g_j(\widetilde{X}^k) \quad \text{as} \quad k \to \infty$$

$$\tau_j^{k+1} = 10^{sign(g_j(\widetilde{X}^k))} \tau_j^k \quad \text{as} \quad k \to \infty$$

For the exterior method, the following updating algorithm may be employed:

$$s_j^{k+1} = 10 s_j^k \quad \text{as} \quad k \to \infty$$

$$r_j^{k+1} = r_j^k - \tau_j^k g_j(\widetilde{X}^k) \quad \text{as} \quad k \to \infty$$

$$\tau_j^{k+1} = 10^{sign(g_j(\widetilde{X}^k))} \tau_j^k \quad \text{as} \quad k \to \infty$$

Once the constrained optimization is transformed to unconstrained optimization, the problem can be solved by using parameterization technique to obtained the global optimum.

3. Parameterization

Although the penalty functions is a continuous nonlinear function, at the optimum point, some penalty functions act similar to discontinuous functions and some act as continuous functions. Also the slope near the optimum point is dramatically changed. This can cause the search becoming violent very easily. Therefore methods to optimize such functions become more difficult. Here, the idea of this approach is based on a damped least square formulation (Mayorga, 2002; Mayorga and Arriaga, 2002) that offers the global optimum.

The technique uses the idea of solving inverse function. Let $\Re^m$ and $\Re^n$ be the *m*- and *n*-dimensional Euclidean spaces and $V = P(X)$ where function $P(X)$ is assumed to be at least twice differentiable.

The concept of this approach is to parameterized variables x and V by a scalar $t \in [t_0, t_f]$ where t_0 and t_f are initial and final values in a finite interval. Let denote $X \equiv X(t)$. The relationship between V and X can be expressed as:

$$V(t) = P(X(t)) \tag{9}$$

Here the problem can be solved according to the idea that variable *V* is set following a desired state transition and try to calculate the corresponding *V*. This establishes an inverse relationship from the above equation. Since the above equation is nonlinear, the analytical inverse relationship cannot be easily obtained. Therefore, this problem can be treated at the inverse functional and rate of change level or indirect fashion.

By differentiating Equation (8), with respect to t, we obtain:

$$\begin{aligned}\dot{V}(t) &= \frac{dV(X(t))}{dt} = \frac{\partial P(X)}{\partial X}\cdot\frac{dX(t)}{dt} \\ &= \nabla P(X(t))\dot{X}(t) \\ &= J(X(t))\dot{X}(t)\end{aligned} \tag{10}$$

where *J*(*X*) is a $1\times n$ Jacobian matrix. Now let $J(X)\equiv J(X(t))$. This allows us to be able to compute $\dot{X}(t)$ as

$$X(t) = J^{+}(X)\dot{V}(t) \tag{11}$$

where $J^{+}(X)$ is a pseudo inverse of the $J(X)$ matrix:

$$J^{+}(X) = J(X)^{T}\left[J(X)J(X)^{T}\right]^{-1} \tag{12}$$

Since function $\dot{V}(t)$ is not known therefore it may be estimated. A technique is able to search for a global optimum was proposed by Mayorga (2002) and Mayorga and Arriaga (2002). The idea is based on the desire to reduce the value of $P(X)$ to estimate the final $P_d(X)$ and proceed until a global optimum is reached. When the search cannot find the lower point, it will *jump* to other valley searching for other optimum. The process continues until the global optimum is found.

$$\dot{V}(t) = \frac{P(X(t_f)) - P(X(t_0))}{t_f - t_0} \tag{13}$$

However, since function $P(X(t_f))$ is not know; the desired $P_d(X(t_f))$ may be approximated as:

$$P_d(X(t_f)) = \alpha P(X(t_0)) \tag{14}$$

where $0\le\alpha\le 1$ is a decreasing ratio for minimization and $\alpha > 1$ for maximization.

Then the slope $\dot{V}(t)$ can be computed by:

$$\dot{V}(t) = \frac{\alpha P(X(t_0)) - P(X(t_0))}{t_f - t_0} \tag{15}$$

Figure 1 shows the parameterization approach performs searching for the global optimum. The search begins by initializing *X*(t_0) and computing P(*X*(t_0)). The desired path and the final point $P(X(t_d))$ are created based on the decreasing ratio.

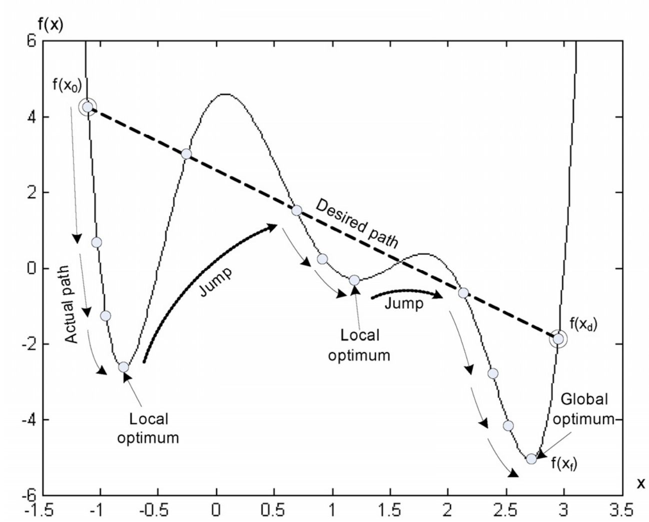

Figure 1. Parameterization approach searching for the global optimum.

The algorithm performing along the actual path by computing function $\dot{X}(t)$ and that function $X(t)$ can be obtained by using an integration method such as the Runge-Kutta. Once a local optimum is obtained, the algorithm then jumps along the path to other valley and continues looking for other local optimum. It jumps once again to other valley looking for another local optimum. Finally the global optimum is obtained.

However, since this approach involves solving the Jacobian and the Hessian matrices and inverse matrix, it is often that the singularity occurs when approaching a local optimum. This will either cause the search being trapped in the local area or the function becomes ill-condition. Therefore the singularity avoidance is concerned (Mayorga, 2000; Mayorga and Sanongboon, 2003).

Here, an exact solution is relaxed and a damped least squares problem is considered (Mayorga, 2000), by adding $[I - J^{+}{}_{WZ\delta}(X)J(X)]v$ to Equation (9) the solution in inexact context can be given by:

$$\dot{X} = J^{+}{}_{WZ\delta}(X)\dot{V} - [I - J^{+}{}_{WZ\delta}(X)J(X)]v \tag{16}$$

where *I* is an $n\times n$ identity matrix, $\delta > 0$ is a small scalar, $J^{+}{}_{WZ\delta}$ is a pseudo inverse with a damping factor on the accuracy to the exact solution, and v is an arbitrary time variation vector of the same dimension as $X(t)$ intended for constraints compliance or singularities avoidance. The pseudo inverse can be obtained from:

$$J^{+}{}_{WZ\delta}(X) = W^{-1}J^{T}(X)\left[J(X)W^{-1}J^{T}(X) + \delta Z^{-1}\right]^{-1} \tag{17}$$

or

$$J^{+}{}_{WZ\delta}(X) = \left[J^{T}(X)ZJ(X) + \delta W\right]^{-1}J^{T}(X)Z \tag{18}$$

where W and Z are the positive definite symmetric matrices that allow invariance to frame reference and scaling. In some cases W is usually diagonal matrix.

The idea of using the nullspace vector v is to maximize the quadratic form of the Hessian matrix in order to satisfy the optimality conditions for the global optimum such that when v is positive definite, the algorithm is in a convex hull of the function. From Equation (17), if $\delta = 0$, the inexact equation be- comes exact solution. The nullspace vector is given by:

$$v_i = -\frac{1}{2}\left(\frac{\partial k(x_i, \dot{x}_i)}{\partial \dot{x}_i}\right) \tag{19}$$

where a scalar $k(x_i, \dot{x}_i)$ is a side criterion that is desired to be optimized. If we suppose that it is desired to decrease mo- notonically a function ϑ over interval $[t_0, t_f]$ by selecting:

$$t(X, \dot{X}) = 2\gamma \frac{\partial \vartheta(X)}{\partial X}\dot{X} \tag{20}$$

where γ is an appropriate scalar [17]. Thus the nullspace vec- tor becomes:

$$v_i = \gamma \frac{\partial s(x_i, x_i)}{\partial x_i} = \gamma \frac{\partial}{\partial x_i}\left(X^T H(X) X\right)^2 \tag{21}$$

where $H(X) \equiv H(X(t))$ is a Hessian matrix of an objective function $P(X)$. The formulation for constrained optimization has been derived:

$$\left[J(X)W^{-1}J^T(X) + \delta Z^{-1}\right]\phi = \dot{V} - J(X)v \tag{22}$$

The above equation involves the dynamic solution of a system in a form of $A(X)\phi = b$ where $A = [J(X)W^{-1}J^T(X) + \delta Z^{-1}]$ and $b = \dot{V} - J(X)v$. Therefore:

$$A(X)\phi = \dot{V} - J(X)v \tag{23}$$

and that function ϕ can be solved by a Gaussian elimination process to obtain $A(X)^{-1}$ such that $\phi = A(X)^{-1}\left[\dot{V} - J(X)v\right]$. According to some math process expressed in [17], $\dot{X}$ can be computed by:

$$\dot{X} = W^{-1}J^T(X)\phi + v \tag{24}$$

For parameters selection, parameters β serves as a sin- gularity avoidance. The large enough β selected will make the matrix $[J(X)J(X)^T]^{-1}$ not becoming ill-conditioned; the small enough β selected will make the problem with singularity avoi- dance still representing the original problem. Therefore, $\beta = \delta Z^{-1} = diag[J(X)J(X)^T]^{-1}$ may be chosen. Matrix W is a diagonal matrix to increase the value of $J(X)W^{-1}J^T(X)$ for the reason as selecting β. Here, $\| J(X)J(X)^T \| \cdot I$ may be selected. For parameter γ, if too large γ is selected, the global minimum may not be obtained. However, $\gamma \leq 1/(X^T H(X) X)^2$ may be utilized. Parameter ξ may be selected in the range of 1 to 5. The smaller value will give more precise solution while large value may lead to ill-conditioned. For selecting α, $\alpha = 0.7 - 0.9999$ is recommended for minimization.

Thus far, the modified penalty functions and parameteri- zation have been expressed. In the next section, an algorithm of the proposed methodology will be illustrated.

4. Algorithm

Refer to Equations (1) and (2), the problem is to mini- mize a nonlinear function subject to constraints. The algori- thm starts with the penalty function process. The desired $P_d(X)$ is estimated from the current penalty function $P_d(X)$. Then com- pute functions $\dot{V}(X)$, $J(X)$, $J^{+}_{WZ\delta}(X)$, $H(X)$, δZ^{-1}, and v. Then compute $\dot{X}(t)$. Finally the Runge-Kutta method is used to obtain $X(t)$. The integration process is terminated when an absolute error of the penalty function $P(X)$ is less than the stopping criterion ε; otherwise the penalty parame- ters are updated according to the global strategy.

Step 0: Set $k = 0$, initialize $X(t)$, s, r, τ and select ξ, α, δ, γ, ε, Z, and W.

Step 1: Compute $P(t)$.

Step 2: Set $t = 0$.

Step 3: Set $t = t_0$ and $X(t) = X_k$.

Step 4: Estimate $P_d(X(t_f)) \approx \alpha P(X(t_0))$.

Step 5: Compute $\dot{V} = [P_d(X(t_f)) - P(X(t_0))]/(t_f - t_0)$.

Step 6: Compute $J(X)$, $J^{+}_{WZ\delta}(X)$, and $H(X)$.

Step 7: Compute δZ^{-1} and null space v. If $X(t)H(X)X(t) \leq \xi$, then δZ^1 and v as shown Equation (21); otherwise $\delta Z^1 = 0$ and $v = 0$.

Step 8: Compute $A(X)$ according to Equation (23).

Step 9: Compute $\phi = A(X)^{-1}[\dot{V} - J(X)v]$.

Step 10: Compute $\dot{X}(t)$ according to Equation (24).

Step 11: Obtain $X_{new}(t)$ using the Runge-Kutta me- thod.

Step 12: Check stopping criterion of the integration loop. If $\left|P(X_{new}(t)) - P(X(t))\right| > \varepsilon$, then $t = t + 1$ and go to step 6; otherwise $X(t_f) = X_{new}(t)$ and go to step 13.

Step 13: Set $\tilde{X}(k) = X(t_f)$.

Step 14: Compute $\tilde{F}(X(k))$.

Step 15: Check stopping criterion of the penalty loop. If $\left|\tilde{F}(X(k)) - F(X(k))\right| > \varepsilon$, then update penalty parameters and set $k = k + 1$ and return to step 1; otherwise $X^* = \tilde{X}(k)$ and $F^* = \tilde{F}$.

5. Implementation

In this section, the methodology is used for solving some problems. This constrained optimization problem is nonlinear and has many local optima. A result obtained from the proposed methodology is compared to ones obtained from the steepest descent, quasi-Newton, the simulated annealing, and the genetic algorithm:

$$\text{Minimize} \quad f(x) = 3(1-x_1)^2 \exp(-x_1^2 - x_2^2) - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right)\cdot \exp(-x_1^2 - x_2^2) - \frac{1}{3}\exp\left(-(x_1+1)^2 - x_2^2\right) + 10 \qquad (25a)$$

$$\text{subject to} \quad g(x) = -2(x_1 - 4) + 3(x_2 - 1) \geq 0 \qquad (25b)$$

The exterior penalty method is chosen as $\Phi(u) = (\max[0, -u])^2$. The starting point is $X = [-0.5, 2.0]$. The parameters are chosen as $\alpha = 0.9$, $\delta = 0.001$, $\sigma = 3$, $\varepsilon = 10^{-3}$, and step size 0f 0.01. The result is shown in the following figure.

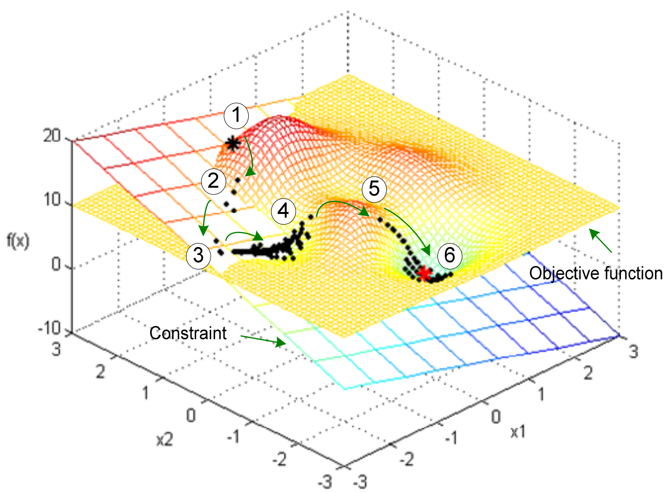

Figure 2. Solving constrained optimization problem using the proposed methodology with singularity avoidance.

As shown in Figure 2, the staring point (point 1) is located near a local optimum and far from the global optimum (point 6). The search then moves to points 2, 3 and 4 very quickly. It moves back and forth in area 4 before jumps out off the valley (point 5) and continues converging to the global optimum. It keeps moving in this area until the stopping criterion is met at $f(x) = 3.7566$.

Table 1. Peak Function with Constraint

Approach	x_1*	x_2*	f*
Steepest descent	–5.1440	8.8174	10.0000
Quasi-Newton	–5.1440	8.8174	10.0000
Simulated annealing	0.2706	–1.4789	3.7566
Genetic algorithm	0.1108	–1.5007	3.8837
Proposed methodology	0.2472	–1.6207	3.4523

The result obtained from the proposed methodology is compared to ones obtained from the steepest descent, quasi-Newton, the simulated annealing, and the genetic algorithm. It shows that the steepest descent and quasi-Newton fail to converge to the global optimum while the proposed methodology converges to the global optimum as the simulated annealing and the genetic algorithm.

6. Least-Cost Treatment of Wastewater

In this section, the proposed method is used for solving the least-cost treatment of wastewater (Chapra and Canale, 2002). As shown in Figure 3, wastewater discharge from big cities are considered a major cause of river pollution. In this paper, a system consists of four cities located on a river and its tributary. Each city generates pollution to the river at a load ing rate P (mg/d). The loading is subject to waste treatment that results in a fractional removal x. The discharge W_i from cities enters the stream and mixes with pollution from upstream sources. The concentration at discharge point is established and then the chemical and biological decomposition processes can remove some pollution out from the river. Note that the headwaters are assumed to be pollution-free.

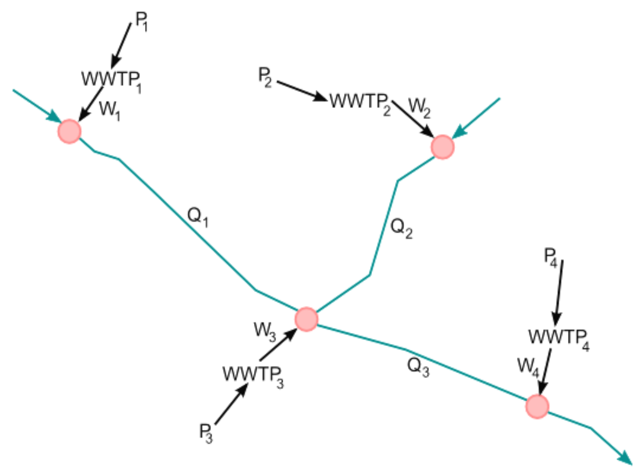

Figure 3. Four wastewater treatment plants discharging pollution to a river system.

The waste treatment costs $1,000/mg removed at each facility. The total daily treatment cost ($1,000/d) can be then estimated as $d_1P_1x_1 + d_2P_2x_2 + d_3P_3x_3 + d_4P_4x_4$. The parameters for the wastewater treatment plants are shown in the following table.

All factors shown above can be combined into the following optimization problem to determine the treatment levels that meet the water quality standard for the minimum cost:

$$\text{Minimize} \quad F(x) = d_1P_1x_1 + d_2P_2x_2 + d_3P_3x_3 + d_4P_4x_4 \qquad (26a)$$

$$\text{Subject to} \quad g_1(x) = cs_1 - \frac{(1-x_1)P_1}{Q_1} \leq 0 \qquad (26b)$$

Table 2. Parameters for Four Wastewater Treatment Plants Discharging Pollution to a River System

City	Pollution loading rate: P_i (mg/d)	Treatment cost: d_i ($\$10^{-6}$/mg)	Concentration: c_i (mg/L)	Flow: Q_i (L/d)	Fractional reduction factor: R	Standard concentration limits: cs (mg/L)
1	1.0×10^9	2	100	1.0×10^7	0.5	20
2	2.0×10^9	2	40	5.0×10^7	3.5	20
3	4.0×10^9	4	47.3	1.1×10^7	0.6	10
4	2.5×10^9	4	22.5	2.5×10^7	-	10

Table 3. Results of Least-Cost Treatment of Wastewater

City	Linear programming		Pattern search GA		Proposed methodology	
	*x**	*c*	*x**	*c*	*x**	*c*
1	0.8000	20.0000	0.8000	20.0000	0.8000	19.9986
2	0.5000	20.0000	0.500	20.0000	0.5000	20.0000
3	0.7500	20.0000	0.7500	20.0000	0.7500	20.0000
4	0.2487	20.0000	0.2487	20.0000	0.2487	20.0000
Total cost	$18,087		$18,087		$18,087	

$$g_2(x) = cs_2 - \frac{(1-x_2)P_2}{Q_2} \le 0 \tag{26c}$$

$$g_3(x) = cs_3 - \frac{R_1Q_1c_1 + R_2Q_2c_2 + (1-x_3)P_3}{Q_3} \le 0 \tag{26d}$$

$$g_4(x) = cs_4 - \frac{R_3Q_3c_3 + (1-x_4)P_4}{Q_4} \le 0 \tag{26e}$$

$$0 \le x_1, x_2, x_3, x_4 \le 1 \tag{26f}$$

The problem was solved using the proposed methodology. The results are shown in the table below. For the proposed approach, the penalty parameters are updated for four times according to the global strategy. The parameters used for this problem are $\alpha = 0.8$, $\delta = 10^{-3}$, $\beta = 10^{-4}$, $\sigma = 5$, $\gamma = 10^{-3}$, $\varepsilon = 10^{-6}$, and step size of 0.1. The minimum is obtained from the proposed methodology is $f(x) = 18{,}087$ which is the same with the results obtained from methods using MATLAB toolboxes including linear program- ming and a pattern search genetic algorithm based technique.

6. Conclusions

A methodology for constrained optimization has been presented. A modified penalty function is used to handle constraints and the parameterization technique is used to optimize the nonlinear function. The improved penalty method eliminate the ill-conditioning and improve the convergence. A transformed unconstrained optimization is then solved for the global optimum. The proposed methodology has been implemented for solving the peak function and least-cost treatment of wastewater. The results show that the proposed methodology is able to search for a global minimum regardless of selecting proper starting point. This is because of the ability to jump out the local valley. Also because of the singularity avoidance, the algorithm will not become ill-conditioned when approaching the local minimum.

The proposed methodology can be improved in certain ways. First, it may be implemented on multi-objective constrained optimization problems. For further improvement, parameter selection may be done to improve quality of the solutions. The proposed methodology requires computing the Jacobian and Hessian matrices to obtain the global optimum. This bottle neck may be eliminated or improved by adopting some techniques such as an artificial neural network may be utilized to reduce computational complexity.

References

Bertsekas, D.P. (1974). *On Penalty and Multiplier Method for Constrained Optimization*, Symposium University of Wisconsin, Madison, Wisconsin.

Chapra, S.C. and Canale, R.P. (2002). *Numerical Method for Engineering*, 4th Edition, McGraw-Hill, New York.

Charalambous, C. (1975). A Method to Overcome the Ill-Conditioning Problem of Differentiable Penalty Functions, *Proceeding of 1975 Midwest Symposium on Circuits & Systems*.

Charalambous, C. (1976). A Negative-Positive Barrier Method for Non-linear Programming, *International Journal of Systems Science*, 7(5), 557-575.

Charalambous, C. (1978). Some Recent Advances in Nonlinear Programming, *Proceeding of the 2nd International Symposium on Large Engineering Systems*, Waterloo, Canada.

Fletcher, R. (1973). An Exact Penalty Function for Nonlinear Programming with Inequalities, *Mathematical Programming*, 5, 129-150, doi:10.1007/BF01580117.

Fonseca, C.M. and Fleming, P.J. (1995). *Multiobjective Optimization and Multiple Constraint Handling with Evolutionary Algorithms II: Application Example, Technical Report*, University of Sheffield, Sheffield, UK.

Haarhof, P.C. and Buys, J.D. (1970). A New Method for Optimization of a Nonlinear Function Subject to Nonlinear Constraints, *The Computer Journal*, 13, 177-179, doi:10.1093/comjnl/13.2.178.

Hestenes, M.R. (1969). Multiplier and Gradient Methods, *J. Optimiz. Theory Appl.*, 4, 303-320, doi:10.1007/BF00927673.

Loughlin, D.H. and Ranjithan, S. (1997). The Neighborhood Constraint Method A Genetic Algorithm-Based Multi-Objective Optimization Technique, *Proceeding of Internal Conference on Ge-*

netic Algorithms, San Mateo, California, pp.666-673.

Mayorga, R.V. (2000). A Metabotics Paradigm for the Intelligent/ Wise Design and Local Motion Planning of Robot Manipulators, *International Symposium of Robotics and Automation*, Monterrey, Mexico.

Mayorga, R.V. (2002). A Paradigm for Continuous Function Optimization: An Intelligent/Wise Systems Implementation, *Proceeding ANIROB/IEEE-RAS International Symposium on Robotics and Automation*, Toluca, Mexico.

Mayorga, R.V. and Arriaga, M. (2002). Multi-Objective Nonlinear Optimization via Parameterization and Inverse Function Approximation: An Artificial Neural Network Implementation, *Proceeding ANIROB/IEEE-RAS International Symposium on Robotics and Automation*, Toluca, Mexico.

Mayorga, R.V. and Quintana, V.H. (1982). Improved Penalty Functions with Applications for Optimal Control Problems, *International Symposium on Mathematical Programming*, Bonn, Germany.

Mayorga, R.V. and Quintana, V.H. (1985). A General Class of Differentiable Penalty Functions for Nonlinear Programming, *12th International Symposium on Mathematical Programming*, M.I.T., Massachusetts, U.S.A.

Mayorga, R.V. and Sanongboon, P. (2003). A Radial Function Approach for Geometrically Bounded Manipulator Inverse Kinematics Computation, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, U.S.A.

Powell, M.J.D. (1969). *A Method for Nonlinear Constraints in Minimization Problems*, Academic Press, pp. 282-298.

Rockafellar, R.T. (1974). Augmented Lagrangian Multiplier Function and Duality in Nonconvex Programming, *SIAM Journal Control*, 12, 268-285.

Surry, P.D., Radeliffe, N.J. and Boyd, I.D. (1995). A Multi-Objective Approach to Constrained Optimization of Gas Supply Networks: The COMOGA Method, Evolutionary Computing, Selected Paper, *Lecture Note in Computer Science*, Springer-Verlag, Sheffield, UK, pp. 166-180.