

Continuous Optimization on Constraint Manifolds

by

Edwin B. Dean
NASA Langley Research Center

Presented at the
TIMS/ORSA Joint National Meeting
Washington D.C.

April 25-27, 1988

Abstract

This paper demonstrates continuous optimization on the differentiable manifold formed by continuous constraint functions. The first order tensor geodesic differential equation is solved on the manifold in both numerical and closed analytic form for simple nonlinear programs. Advantages and disadvantages with respect to conventional optimization techniques are discussed.

Introduction

For many years operations research practitioners have sporadically considered using differential equations for optimization. Tanabe [1,2] took this approach into the realm of modern mathematics by deriving the properties of the tensor first order differential equation which can be used for optimization. This paper uses his theoretical foundations, extends his equation to a geodesic form, and demonstrates that nonlinear optimization problems can be solved both numerically and analytically with differential equations on the continuous differentiable manifolds formed by constraining a vector of continuous functions.

To be consistent with the illuminating tensor notation of Gerretsen [3], row vector notation is used throughout this paper. Thus, a vector v with components a_i in basis $\{b_1, \dots, b_n\}$ may be expressed as

$$v = a B$$

where

$$a = [a_1 \dots a_n] \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ \dots \\ b_n \end{bmatrix}$$

All computer solutions were performed on a Macintosh Plus TM using Lightspeed Pascal TM.

Summary of Theory

The following theory is based on concepts from differentiable manifolds and differential geometry as described by Thorpe [4] and Boothby [5].

An n-dimensional manifold is a connected, locally compact space with a countable basis, each point of which has a neighborhood homeomorphic to euclidian n-space.

A C^k differentiable manifold is a manifold with additional mathematical properties imposed which permit the definition of compatible coordinate systems on the manifold which are mapped by diffeomorphic functions from the manifold into euclidian n-space.

This mapping is depicted by Figure 1.

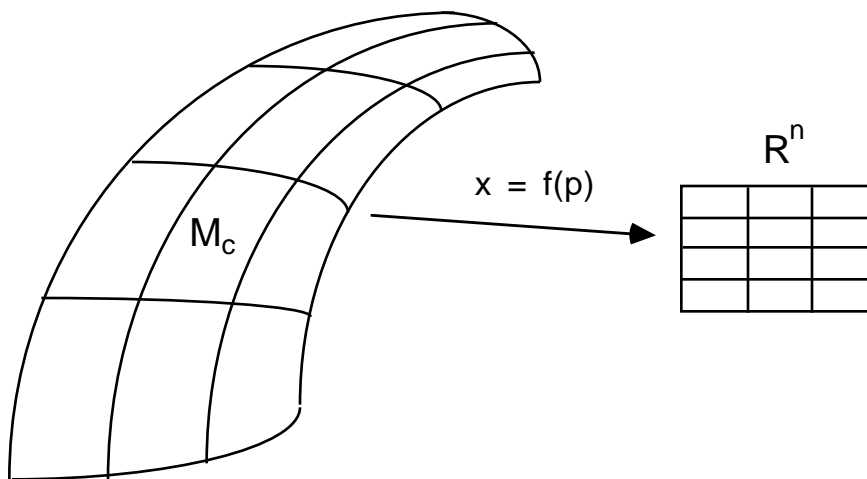


Figure 1
Manifold Patch with Mapping into R^n .

$M_c = \{ p \mid g(p) = c \}$ can be shown to be a C^k differentiable manifold.

Such a manifold may be called a "constraint manifold" since it is totally defined given a set of C^k constraint functions.

The nonlinear programming problem may be stated as

extremalize $f(x)$
 over x
 subject to $g(x) = \text{constant}$.

For C^k functions f and g with $k > 0$ and for the Jacobian matrix

$$\partial_x g = \begin{bmatrix} \partial_1 g_1 & \dots & \partial_n g_1 \\ \dots & \dots & \dots \\ \partial_1 g_m & \dots & \partial_n g_m \end{bmatrix}$$

where $\partial_j g_i = \frac{\partial g_i}{\partial x_j}$,

there exists tensors N_p and T_p at the point p defined by

$$N_p = \partial_p^T g (\partial_p g \partial_p^T g)^{-1} \partial_p g \quad \text{and} \quad T_p = I - N_p$$

where $\partial_p^T g$ is the transpose of $\partial_p g$ and the subscript p represents evaluation of the functions at p .

Operations research analysts will recognize T_p as a generalization of the projection matrix in the Rosen gradient projection algorithm [6] to evaluation at the tensor operating point p .

Geometrically, T_p projects onto the tangent space to the constraint manifold at p , and N_p projects onto the normal space to the constraint manifold at p .

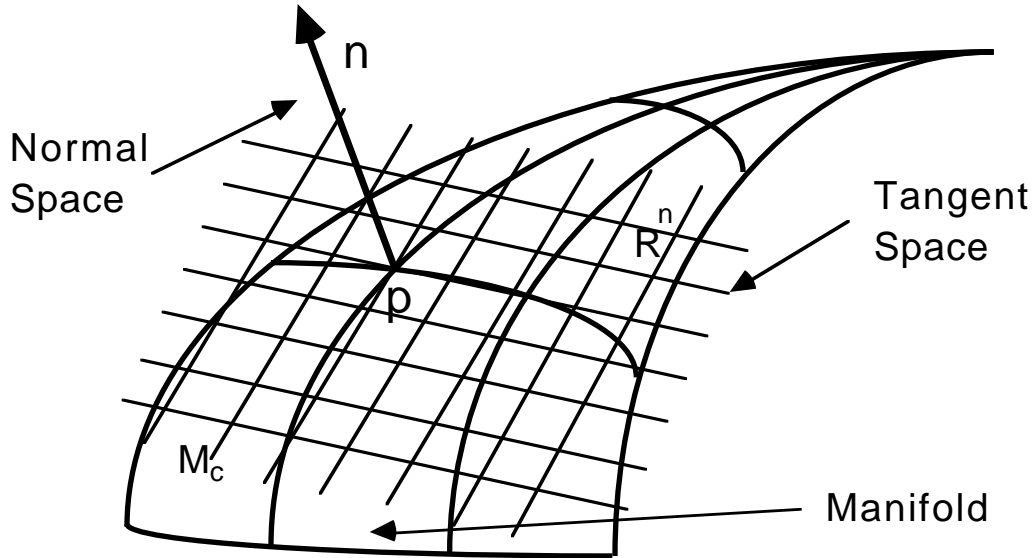


Figure 2
Normal and Tangent Spaces to the Manifold
at the Tensor Operating Point p.

A manifold M_c is covered by a vector field if and only if at each point p there is a vector $v(p)$. If $f(x)$ is C^k function over M_c then the gradient $\partial_p f$ forms a C^k vector field over M_c . If $f(x)$ and $g(x)$ are C^k functions over M_c then $\partial_p f T_p$ forms a C^k vector field over M_c .

It is important to note that $\partial_p f T_p$ is the restriction of $\partial_p f$ to M_c . That means that $\partial_p f T_p$ is in the tangent space to M_c with origin at p .

Tanabe [2] has shown that under appropriate second order conditions, the flow

$$\dot{p} = \pm \partial_p f T_p$$

extremalizes $f(x)$ on M_c by converging to a local extremum at steady state.

A flow is a geodesic on a manifold M_c if and only if all acceleration is normal to the manifold.

Consider the unit velocity flow

$$\dot{p} = \frac{\pm \partial_p f T_p}{\sqrt{\partial_p f T_p \partial_p^T f}}$$

on M_c , then

- (a) $\dot{p} \dot{p}^T = 1$,
- (b) $\ddot{p} = 0$,
- (c) any acceleration must be normal to M_c ,
- (d) the unit velocity flow is a geodesic on M_c ,
- (e) the unit velocity flow is the shortest distance on M_c between any initial condition on M_c and an extremum of M_c , and
- (f) the distance between the initial condition and an extremum is the arc length of the path and hence a straight line on M_c .

A trivial extension of Tanabe's results [2] shows that under appropriate second order conditions the unit velocity flow

$$\dot{p} = \frac{\pm \partial_p f T_p}{\sqrt{\partial_p f T_p \partial_p^T f}}$$

extremalizes $f(x)$ on M_c by converging to a local extremum at steady state.

At the extrema

- (a) $\dot{p} = 0$,
- (b) $\partial_p f = \partial_p f N_p \Rightarrow$ All objective imparted velocity is normal to M_c .

- (c) If no two $\partial_p g_i$ are colinear at p
- (1) N_p has m eigenvalues $\lambda_i(p)$ such that $\lambda_i(p) = 1$,
 - (2) $\partial_p f$ can be expressed in the m dimensional basis of the eigenvectors associated with the m unit eigenvalues,
 - (3) N_p has n-m eigenvalues $\lambda_i(p)$ such that $\lambda_i(p) = 0$, and
- (d) Nonlinear equations exist, which, if solvable, lead to a closed form solution.

The Lagrange multipliers μ_i are

$$\mu(p) = \partial_p f \partial_p^T g (\partial_p g \partial_p^T g)^{-1}$$

which gives

$$\dot{p} = \partial_p f - \mu(p) \partial_p g.$$

At any point of the trajectory, the Lagrange multipliers μ_i are the normal components of the velocity which have been subtracted from the velocity imparted by the objective $f(x)$ in order to restrict the trajectory to M_c .

Circular Manifold Analytical Solution

Simple problems can be solved analytically for the dynamic path or just for the optimum points with this approach. The circular constraint manifold is solved below for a linear objective function to obtain the optimum point \hat{p} and optimum objective value $f(\hat{p})$. The problem is stated as

$$\begin{array}{ll} \text{maximize} & f(x) = a_1 x_1 + a_2 x_2 \\ \text{over } x_1 \text{ and } x_2 & \\ \text{subject to} & g(x) = x_1^2 + x_2^2 = r^2. \end{array}$$

Evaluated at the point p we have

$$\partial_p g = [2p_1 \ 2p_2]$$

$$\partial_p f = [a_1 \ a_2]$$

$$N_p = \partial_p^T g (\partial_p g \partial_p^T g)^{-1} \partial_p g = \frac{1}{r^2} \begin{bmatrix} p_1^2 p_1 p_2 \\ p_2 p_1 p_2^2 \end{bmatrix}$$

$$T_p = I - N_p = \frac{1}{r^2} \begin{bmatrix} p_2^2 & -p_1 p_2 \\ -p_2 p_1 & p_1^2 \end{bmatrix} \text{ and}$$

$$\partial_p f T_p = \frac{1}{r^2} [a_1 p_2^2 - a_2 p_1 p_2 - a_1 p_2 p_1 + a_2 p_1^2]$$

Setting

$$\partial_p f T_p = 0$$

we have

$$a_1 \dot{p}_2 - a_2 \dot{p}_1 = 0.$$

The constraint provides

$$\dot{p}_1^2 + \dot{p}_2^2 = r^2.$$

Combining the two equations we have

$$\dot{p}_1 = \frac{a_2 r}{\sqrt{a_1^2 + a_2^2}},$$

$$\dot{p}_2 = \frac{a_1 r}{\sqrt{a_1^2 + a_2^2}}, \text{ and}$$

$$f(\dot{p}) = \frac{2a_1 a_2 r}{\sqrt{a_1^2 + a_2^2}}.$$

Circular Manifold Numerical Solution

The steepest descent trajectory on the circular constraint manifold is found by solving

$$\begin{aligned}\dot{p} &= \partial_p f T_p \\ &= \frac{1}{r^2} [a_1 p_2^2 - a_2 p_1 p_2 - a_1 p_2 p_1 + a_2 p_1^2]\end{aligned}$$

The velocity along the manifold is

$$\begin{aligned}v(p) &= (\dot{p} \dot{p}^T)^{1/2} \\ &= (\partial_p f T_p T_p^T \partial_p f^T)^{1/2} \\ &= (\partial_p f T_p \partial_p f^T)^{1/2}.\end{aligned}$$

For an n dimensional sphere and hyperplane the velocity can be shown to be

$$v(p) = \left(a a^T - \frac{(a p^T)^2}{r^2} \right)^{1/2}.$$

Using a fourth order Runge Kutta algorithm from Flanders [7] extended to multidimensional systems we have the following numerical trajectory. The radius has been set at 1, the components a_i set to 1, t is the trajectory measure, r^2 is the square of the radius of the sphere, v is the trajectory velocity at $p(t)$, and p_1, p_2 are the components of $p(t)$.

Note that the radius stays very close to 1.000000 throughout the trajectory. Consequently the algorithm follows the manifold closely. Also, since the unit velocity trajectory is not taken above, the velocity drops to zero as the optimum value is approached. This is actually desirable with the algorithm used. With a unit velocity trajectory it is difficult near the optimum to converge to the optimum because the constant distance step oversteps the optimum and hangs up.

Table 1
Steepest Descent of a Linear Objective

on a Circular Constraint Manifold of Radius One

t	r ²	v	p ₁	p ₂
0.00	1.000000	1.400000	-0.800000	0.600000
0.40	0.999792	1.296556	-0.365864	0.930556
0.80	0.999000	0.924766	0.072559	0.996863
1.20	0.998718	0.572417	0.360155	0.932205
1.60	0.998855	0.334783	0.519319	0.853910
2.00	0.999054	0.191984	0.604283	0.796176
2.40	0.999294	0.109368	0.650076	0.759405
2.80	0.999437	0.062178	0.675144	0.737304
3.20	0.999584	0.035324	0.689081	0.724397
3.60	0.999678	0.020064	0.696892	0.716952
4.00	0.999755	0.011395	0.701300	0.712694
4.40	0.999816	0.006472	0.703799	0.710270
4.80	0.999851	0.003676	0.705214	0.708889
5.20	0.999890	0.002087	0.706023	0.708111
5.60	0.999912	0.001185	0.706483	0.707668
6.00	0.999930	0.000673	0.706745	0.707418
6.40	0.999944	0.000382	0.706896	0.707278
6.80	0.999952	0.000217	0.706981	0.707198
7.20	0.999956	0.000123	0.707030	0.707153
7.60	0.999959	0.000070	0.707057	0.707127
8.00	0.999960	0.000040	0.707073	0.707113
8.40	0.999961	0.000022	0.707082	0.707104
8.80	0.999962	0.000013	0.707087	0.707100
9.20	0.999962	0.000007	0.707090	0.707097
9.60	0.999962	0.000004	0.707091	0.707095
10.00	0.999962	0.000002	0.707092	0.707095

The unit velocity geodesic trajectory on the circular constraint manifold is found by solving

$$\dot{p} = \frac{\partial_p f T_p}{v(p)}$$

Spherical Manifold Numerical Solution

The projection equations for the spherical constraint manifold are derived below.

The problem is stated as

$$\begin{aligned} &\text{maximize} && f(x) = a_1 x_1 + a_2 x_2 + a_3 x_3 \\ &\text{over } x_1, x_2 \text{ and } x_3 \\ &\text{subject to} && g(x) = x_1^2 + x_2^2 + x_3^2 = r^2. \end{aligned}$$

Evaluated at the point p we have

$$\partial_p g = [2p_1 \ 2p_2 \ 2p_3]$$

$$\partial_p f = [a_1 \ a_2 \ a_3]$$

$$N_p = \partial_p^T f (\partial_p g \ \partial_p^T g)^{-1} \partial_p g = \frac{1}{r^2} \begin{bmatrix} p_1^2 & p_1 p_2 & p_1 p_3 \\ p_2 p_1 & p_2^2 & p_2 p_3 \\ p_3 p_1 & p_3 p_2 & p_3^2 \end{bmatrix}$$

$$T_p = I - N_p = \frac{1}{r^2} \begin{bmatrix} p_2^2 + p_3^2 - p_1 p_2 & -p_1 p_3 \\ -p_2 p_1 & p_1^2 + p_3^2 - p_2 p_3 \\ -p_3 p_1 & -p_3 p_2 & p_1^2 + p_2^2 \end{bmatrix} \text{ and}$$

$$\begin{aligned} \partial_p f T_p = \frac{1}{r^2} [&a_1(p_2^2 + p_3^2) - a_2 p_1 p_2 - a_3 p_1 p_3, \\ &-a_1 p_2 p_1 + a_2(p_1^2 + p_3^2) - a_3 p_2 p_3, \\ &-a_1 p_3 p_1 - a_2 p_3 p_2 + a_3(p_1^2 + p_2^2)] \end{aligned}$$

The steepest descent trajectory on the spherical constraint manifold is found by solving

$$\dot{p} = -\partial_p f T_p$$

Using a fourth order Runge Kutta algorithm from Flanders [7] extended to multidimensional systems we have the following numerical trajectory. The radius has been set at 1, the components a_i set to 1, t is the trajectory measure, r^2 is the square of the radius of the sphere, v is the trajectory velocity at $p(t)$, and $p_1, p_2, \text{ and } p_3$ are the components of $p(t)$.

Table 2
Steepest Descent of a Linear Objective
on a Spherical Constraint Manifold of Radius One

t	r ²	v	p ₁	p ₂	p ₃
0.00	1.000000	1.720465	-0.800000	0.600000	0.000000
0.40	0.999827	1.479468	-0.330427	0.873341	0.357660
0.80	0.998985	0.902703	0.107880	0.842055	0.527532
1.20	0.998620	0.477765	0.351083	0.739594	0.573028
1.60	0.998678	0.242504	0.468021	0.665234	0.580602
2.00	0.998835	0.121738	0.523763	0.622780	0.580217
2.40	0.998945	0.060952	0.550750	0.600333	0.578982
2.80	0.999054	0.030496	0.564019	0.588830	0.578114
3.20	0.999116	0.015256	0.570590	0.583004	0.577624
3.60	0.999173	0.007631	0.573867	0.580078	0.577372
4.00	0.999201	0.003817	0.575500	0.578608	0.577246
4.40	0.999230	0.001909	0.576321	0.577875	0.577187
4.80	0.999249	0.000954	0.576732	0.577509	0.577159
5.20	0.999263	0.000477	0.576938	0.577326	0.577148
5.60	0.999270	0.000238	0.577040	0.577234	0.577144
6.00	0.999273	0.000119	0.577091	0.577188	0.577142
6.40	0.999275	0.000059	0.577116	0.577165	0.577141
6.80	0.999275	0.000030	0.577129	0.577153	0.577141
7.20	0.999276	0.000015	0.577135	0.577147	0.577141
7.60	0.999276	0.000007	0.577138	0.577144	0.577141
8.00	0.999276	0.000004	0.577140	0.577143	0.577141
8.40	0.999276	0.000002	0.577141	0.577142	0.577141
8.80	0.999276	0.000001	0.577141	0.577142	0.577141
9.20	0.999276	0.000000	0.577141	0.577141	0.577141

Since the unit velocity trajectory is not taken above, the velocity drops toward zero as the optimum value is approached.

The unit velocity geodesic trajectory on the spherical constraint manifold is found by solving

$$\dot{p} = \frac{\partial_p f T_p}{v(p)}$$

Table 3 illustrates the same solution as in Table 2 except the trajectory is along the unit velocity geodesic. A constant trajectory parameter step fourth order multidimensional Runga Kutta algorithm from Chua and Lin [8] was used. The trajectory parameter here was set at .01 and only every tenth step is shown until near the optimum value. After t=1.6 every step is shown. The optimum is between t=1.68 and t=1.69, but the algorithm failed because there is no

variable step end game to permit convergence to the optimum. Note the improved tracking of the manifold surface with $r^2 \geq 0.999999$. Note also that the trajectory parameter, in this case the arc length or minimum distance trajectory, reaches the optimum at a much smaller value corresponding to the shorter distance along the trajectory.

Table 3
Geodesic Descent of a Linear Objective
on a Spherical Constraint Manifold of Radius One

t	r ²	p ₁	p ₂	p ₃
0.00	1.000000	-0.800000	0.600000	0.000000
0.10	1.000000	-0.747261	0.661993	0.058027
0.20	1.000000	-0.687055	0.717371	0.115474
0.30	1.000000	-0.619984	0.765582	0.171768
0.40	1.000000	-0.546719	0.806143	0.226345
0.50	1.000000	-0.467991	0.838649	0.278660
0.60	1.000000	-0.384587	0.862776	0.328192
0.70	1.000000	-0.297341	0.878282	0.374444
0.80	1.000000	-0.207123	0.885013	0.416955
0.90	1.000000	-0.114836	0.882901	0.455299
1.00	1.000000	-0.021402	0.871968	0.489095
1.20	0.999999	0.165172	0.824160	0.541736
1.30	0.999999	0.256448	0.787763	0.560056
1.40	0.999999	0.345161	0.743495	0.572781
1.50	0.999999	0.430426	0.691799	0.579782
1.60	0.999999	0.511388	0.633192	0.580990
1.61	0.999999	0.519215	0.626973	0.580791
1.62	0.999999	0.526991	0.620692	0.580535
1.63	0.999999	0.534714	0.614349	0.580220
1.64	0.999999	0.542383	0.607945	0.579847
1.65	0.999999	0.549996	0.601481	0.579416
1.66	0.999999	0.557552	0.594960	0.578928
1.67	0.999999	0.565043	0.588387	0.578382
1.68	0.999999	0.572355	0.581866	0.577790

Conclusions

From the viewpoint of continuous optimization on constraint manifolds the theory and application of the nonlinear program reduces to the study of the zeros and the geodesic flows of the vector field

$$\frac{\partial_p f T_p}{\sqrt{\partial_p f T_p \partial_p^T f}}$$

The theory and examples in this paper provide the jumping off point for many additional studies. Both numerical and analytical dynamic solutions should provide valuable information about the geometry of the constraint manifolds by studying the natural straight lines on the manifold and the sensitivity of these geodesics to initial conditions. Many other possibilities exist.

Preliminary numerical solutions indicate that without special tailoring of the differential equation solution algorithms, conventional minimum seeking optimization algorithms are faster. However, special algorithmic end games may reduce the advantage of conventional algorithms substantially.

On the other hand, the amount of geometrical understanding of the constraint manifolds provided by the geodesic flows is far superior.

Finally, as opposed to conventional nonlinear programming algorithmic approaches, the geodesic flow approach is in a mathematical form readily analyzed by recent advances in global dynamic theory.

References

- [1] Tanabe, K., "Algorithm for the Constrained Maximization Technique for Nonlinear Programming," Proceedings of the Second Hawaii International Conference on Systems Science , 1969, 487-490.
- [2] Tanabe, K., "An Algorithm for the Constrained Maximization in Nonlinear Programming," J. Operations Research Soc. of Japan , Vol. 17, No. 4, December 1974.
- [3] Gerretsen, J. C. H., Lectures on Tensor Calculus and Differential Geometry , P. Noordhoff, Groningen, The Netherlands, 1962.
- [4] Thorpe, J. A., Elementary Topics in Differential Geometry , Springer-Verlag, New York, NY, 1979.
- [5] Boothby, W. M., An Introduction to Differentiable Manifolds and Riemannian Geometry , Academic Press, New York, NY, 1975.
- [6] Rosen, J. B., "The Gradient Projection Method for Nonlinear Programming: Part II, Nonlinear Constraints," J. Soc. Industr. Appl. Math. , 9, 4 (1961).
- [7] Flanders, H., Scientific Pascal , Reston Publishing Company, Reston, VA, 1984.
- [8] Chua, L. O. and Lin, P., Computer-Aided Analysis of Electronic Circuits, Algorithms and Computational Techniques , Prentice-Hall, Englewood Cliffs, NJ, 1975.