SLAC-PUB-1292
(M)
August 1973

NON-LINEAR CONSTRAINT ELIMINATION IN HIGH DIMENSIONALITY THROUGH REVERSIBLE TRANSFORMATIONS*

J.H. Friedman and S. Steppel

Stanford Linear Accelerator Center
Stanford, California

June 1973

ABSTRACT

A method for accommodating a large class of equality constraints that appear in multidimensional constrained optimization and sampling problems is discussed. A recursive procedure is developed for reducing the constraints on the full n-dimensional space to similar constraints on progressively lower dimensional subspaces. The technique can be applied to multiply constrained problems as long as each constraint can be cast in the general form

$$F\left[\sum_{i=1}^{n} f_i(X_i)\right] = \text{constant}$$

where $\{f_i\}_{i=1}^{n}$ are arbitrary functions of the individual coordinates and F is an invertible function.

(Submitted to The Computer Journal)

*Work supported by the U.S. Atomic Energy Commission.

INTRODUCTION

In many applications one is forced to deal with functions of many variables where there exist non-linear relations among the variables. Often this problem can be formulated in terms of a scalar function defined over a multidimensional space, with the constraints defining a lower dimensional manifold embedded in the space. Frequently one is faced with optimizing or integrating the function over the manifold.

The classical approach to the optimization problem with equality constraints is the method of Lagrange multipliers. For the case of m constraints in an n-dimensional space, this technique converts the problem to one of solving a set of m + n non-linear equations. The difficulties involved in finding solutions to sets of simultaneous non-linear equations limits the applicability of this approach.[1] Another approach is to reduce the constrained optimization problem to a sequence of unconstrained problems by the incorporation of penalty functions into the objective function.[2] A third method is to transform the independent variables to coordinates that automatically satisfy the constraints, leaving the objective function unchanged. When such transformations can be found, this is generally preferable to the first two methods. The transformation approach reduces the dimensionality of the problem to n - m, rather than increasing it to n + m, as is the case with Lagrange multipliers. Furthermore, only one unconstrained optimization is necessary rather than a series of them, as required by the incorporation of penalty functions. Generally, these advantages result in increased stability and computational efficiency. Very few generally applicable transformation techniques have been developed, however, due to the difficulty of finding the appropriate transformations for most problems.

For numerical integration over manifolds embedded in higher dimensional spaces, one must choose function evaluation points that lie on the manifold. In general this is difficult, especially when a uniform distribution of sample points is required. When possible, one can perform a transformation to the manifold from the unit hypercube of corresponding dimension. For certain applications, there have also been discovered specialized tricks* for random sampling on manifolds embedded in higher dimensional spaces.

This report describes a procedure for obtaining reversible transformations from a large class of manifolds defined by one or more equality constraints to the corresponding Euclidean space. This method utilizes a recursive procedure that converts the constraints on the full n-dimensional space to similar constraints on progressively lower dimensional subspaces. The problem is then reduced to applying the constraints on these low dimensional subspaces, where they may often be dealt with directly.

---

*See reference 3 for a large collection of these specialized techniques.

Single Equality Constraint

Consider a constrained n-dimensional integration of the following form

$$\int dV_n(C_n) = \int \left(\prod_{i=1}^{n} dX_i\right) \delta \left\{C_n - F\left[\sum_{i=1}^{n} f_i(X_i)\right]\right\}, \tag{1}$$

where the $f_i(X_i)$ are arbitrary functions of the individual coordinates and F is an invertible function. This is an integration over an n-1 dimensional manifold defined by the equality constraint in the argument of the Dirac δ-function. Let

$$C_k = F\left[\sum_{i=1}^{k} f_i(X_i)\right] \qquad (k<n) \tag{2a}$$

and

$$C_{n-k} = F\left[\sum_{i=k+1}^{n} f_i(X_i)\right] \tag{2b}$$

then the RHS of eqn. 1 can be re-expressed as

$$\int \delta\left\{C_n - F\left[F^{-1}(C_k) + F^{-1}(C_{n-k})\right]\right\} \delta\left\{C_k - F\left[\sum_{i=1}^{k} f_i(X_i)\right]\right\}$$

$$\times\, \delta\left\{C_{n-k} - F\left[\sum_{i=k+1}^{n} f_i(X_i)\right]\right\} dC_k\, dC_{n-k} \prod_{i=1}^{n} dX_i$$

or

$$\int dV_n(C_n) = \int dC_k dC_{n-k} \quad dV_k(C_k)\, dV_{n-k}(C_{n-k})$$

$$\times\, \delta\left\{C_n - F\left[F^{-1}(C_k) + F^{-1}(C_{n-k})\right]\right\} \tag{3}$$

where $dV_k(C_k)$ and $dV_{n-k}(C_{n-k})$ are defined by eqn. (1).

The constrained n-dimensional integration of eqn. 1 has been converted in eqn. 3 to two similarly constrained integrations on k and n-k dimensions, and a constrained two-dimensional integration. The procedure that led from eqn. 1 to eqn. 3 can now be analogously applied to $dV_k(C_k)$ and $dV_{n-k}(C_{n-k})$. This yields a further reduction of dimensionality for the constrained integrations and the addition of two more two-dimensional constrained integrations. This dimension reducing process can be continued until the constrained integrations are of sufficiently low dimensionality to be handled directly. Thus eqn. 3 provides a framework for deriving recursive procedures that utilize known solutions in lower dimensional cases to solve higher dimensional problems.

There is considerable flexibility in the degree by which the dimensionality can be reduced at each step, namely, any k in the range $1 \leq k \leq n-1$ may be used.

The number of these dimension reducing steps, and thus the ultimate reduction in the dimensionality of the constrained integration, is also arbitrary. However, it is clear from eqn. 3 that the largest reduction that can be achieved is a series of n-1 two-dimensional constrained integrations.

We now illustrate how this recursive strategy can be implemented to solve specific problems.

A. <u>Solid Angle Transform (SAT)</u>

This transformation maps the n-1 dimensional unit hypercube to the surface of a sphere in n-dimensions. Such a surface can be defined by the equality constraint

$$\sqrt{\sum_{i=1}^{n} X_i^2} = r_n \tag{4}$$

This constraint occurs in n-dimensional random direction sampling as well as many constrained optimization problems.

For this constraint, eqn. 1 becomes

$$\int dV_n(r_n) = \int \prod_{i=1}^{n} dX_i \quad \delta(r_n - \sqrt{\sum_{i=1}^{n} X_i^2}\ ). \tag{5}$$

Applying eqn. 3 with k=2

$$\int dV_n(r_n) = \int dr_2 dr_{n-2}\ dV_2(r_2)\ dV_{n-2}(r_{n-2}) \cdot$$
$$x\ \delta(r_n - \sqrt{r_2^2 + r_{n-2}^2}\ ).$$

Converting r_2 and r_{n-2} to polar coordinates

$$r_2 = \zeta \cos\theta,\ r_{n-2} = \zeta \sin\theta$$

and integrating over ζ, we obtain

$$\int dV_n(r_n) = r_n \int_0^{\pi/2} d\theta\ dV_2(r_n \cos\theta)\ dV_{n-2}(r_n \sin\theta). \tag{6}$$

Now

$$dV_2(r_n \cos\theta) = \int dX_1 dX_2\ \delta(r_n \cos\theta - \sqrt{X_1^2 + X_2^2}\ )$$

Again using polar coordinates

$$X_1 = R\cos\varphi \qquad X_2 = R\sin\varphi$$

and integrating over R

$$dV_2\ (r_n \cos\theta) = r_n \cos\theta \int_0^{2\pi} d\varphi = 2\pi r_n \cos\theta \int_0^1 d\mu_1$$

Substituting this into eqn. (6)

$$\int dV_n(r_n) = 2\pi r_n^2 \int_0^1 d\mu_1 \int_0^{\pi/2} \cos\theta d\theta \, dV_{n-2}(r_n \sin\theta)$$

$$= 2\pi r_n^2 \int_0^1 d\mu_1 \int_0^1 d\mu_2 \, dV_{n-2}(r_n \mu_2) \tag{7}$$

Note that

$$\begin{aligned} X_1 &= r_n \cos(\sin^{-1}\mu_2) \cos(2\pi\mu_1) \\ X_2 &= r_n \cos(\sin^{-1}\mu_2) \sin(2\pi\mu_1) \; . \end{aligned} \tag{8}$$

Reapplying eqn. 3 to dV_{n-2} in eqn. 7 and proceeding as above

$$\int dV_n(r_n) = (2\pi)^2 r_n^4 \int_0^1 d\mu_1 \int_0^1 \mu_2^2 d\mu_2 \int_0^1 d\mu_3 \int_0^1 d\mu_4 \, dV_{n-4}(r_n\mu_2\mu_4) \tag{9}$$

where

$$\begin{aligned} X_3 &= r_n\mu_2 \cos(\sin^{-1}\mu_4) \cos(2\pi\mu_3) \\ X_4 &= r_n\mu_2 \cos(\sin^{-1}\mu_4) \sin(2\pi\mu_3). \end{aligned} \tag{10}$$

Continuing this two at a time reduction m=[n/2-1]* times one obtains

$$(2\pi)^m r_n^{2m} \left[\prod_{i=1}^{m} \int_0^1 d\mu_{2i-1}\right] \int_0^1 \prod_{i=1}^{m} d\mu_{2i} \, \mu_{2i}^{2(m-i)} \tag{11}$$

$$\times \, dV_{n-2m}(r_n \prod_{i=1}^{m} \mu_{2i}) \; .$$

*[f] is the greatest integer less than or equal to f.

If n is even, then n-2m = 2. In this case

$$dV_{n-2m} = 2\pi r_n \prod_{i=1}^{m} \mu_{2i} \int_0^1 d\mu_{n-1}$$

so that eqn. 11 becomes

$$(2\pi)^{n/2}\, r_n^{n-1} \left[\prod_{i=1}^{n/2} \int_0^1 d\mu_{2i-1}\right] \int_0^1 \prod_{i=1}^{(n-2)/2} d\mu_{2i}\, \mu_{2i}^{n-2i-1} \tag{12}$$

or

$$\frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}\; r_n^{n-1} \prod_{i=1}^{n-1} \int_0^1 d\eta_i \tag{13}$$

where

$$\eta_{2i-1} = \mu_{2i-1} \qquad 1 \leq i \leq n/2$$

$$\eta_{2i} = \mu_{2i}^{n-2i} \qquad 1 \leq i \leq n/2-1 \; ,$$

with

$$X_{2i} = \left[r_n \prod_{j=1}^{i-1} \eta_{2j}^{1/(n-2j)}\right] \cos\left[\sin^{-1}\eta_{2i}^{1/(n-2i)}\right] \sin(2\pi\eta_{2i-1}) \qquad (1 \leq i \leq n/2-1)$$

$$X_n = \left[r_n \prod_{j=1}^{n/2-1} \eta_{2j}^{1/(n-2j)}\right] \sin(2\pi\eta_{n-1}) \tag{14}$$

$$X_{2i-1} = X_{2i} \cot(2\pi\eta_{2i-1}) \qquad (1 \leq i \leq n/2)$$

If n is odd, then n-2m = 3. In this case,

$$dV_{n-2m} = 4\pi r_n^2 \prod_{i=1}^{m} \mu_{2i}^2 \int_0^1 d\mu_{n-2} \int_0^1 d\mu_{n-1} \; .$$

Inserting this into eqn. 11 and following the procedures used for the even case, we once again obtain eqn. 13, with the transformation

$$X_n = \left[r_n \prod_{j=1}^{\frac{n-3}{2}} \eta_{2j}^{1/(n-2j)} \right] (2\eta_{n-2} - 1)$$

$$X_{n-1} = \left[r_n \prod_{j=1}^{\frac{n-3}{2}} \eta_{2j}^{1/(n-2j)} \right] (\eta_{n-2} - \eta_{n-2}^2)^{1/2} \sin(2\pi\eta_{n-1})$$

$$X_{2i} = \left[r_n \prod_{j=1}^{i-1} \eta_{2j}^{1/(n-2j)} \right] \cos \left[\sin^{-1} \eta_{2i}^{1/(n-2i)} \right] \sin(2\pi\eta_{2i-1}) \qquad 1 \le i \le (n-3)/2 \tag{15}$$

$$X_{2i-1} = X_{2i} \cot (2\pi\eta_{2i-1}) \qquad 1 \le i \le (n-1)/2 \; .$$

Equations 14 and 15 represent a transformation from the (n-1) dimensional unit hypercube to the surface of an n-dimensional sphere of radius r_n. The η_i $(1 \le i \le n-1)$ are the hypercube variables, while the X_i $(1 \le i \le n)$ are the constrained variables on the n-dimensional space. From eqn. 13 we see that the Jacobian of this transformation,

$$J_n(r_n) = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})} r_n^{n-1} \quad ,$$

in a constant, namely the well known expression for the surface area of an n-dimensional sphere of radius r_n. It is a volume preserving one to one onto transformation. The inverse transformation can easily be obtained by solving equations 14 or 15 for the η's in terms of the X-coordinates.

In deriving this transformation, we applied eqn. 3 to our constraint (eqn. 4) using k=2. As discussed above, this choice of k is not unique.

Choosing k=1 and following the same procedure would have led to the standard n-dimensional spherical polar coordinate representation[4] as the transformation. This transformation has the Jacobian

$$J_n(r_n) = r_n^{n-1} \prod_{i=1}^{n-2} \sin^{n-1-i} \varphi_i \, .$$

In order to obtain a constant Jacobian, the following additional transformation must be performed

$$\frac{\int_0^{\varphi_i} \sin^{n-i-1} \varphi \, d\varphi}{B(\frac{n-i}{2}, 1/2)} = \eta_i \qquad (1 \le i \le n-2) \tag{17}$$

$$\theta = 2\pi \, \eta_{n-1}$$

Thus using k=1 for this problem has the advantage that we need not consider special cases for even and odd dimensionality. If volume preservation is desired, however, eqn. 17 must be solved numerically for φ_i in terms of η_i.

Choosing different values of k for this problem corresponds to the decomposition of the n-dimensional sphere into lower-dimensional spheres in various ways, as discussed by Shelupsky.[5]

B. Power Product Transform (PPT)

For a second example of a single equality constraint, we consider one of the form

$$\prod_{i=1}^{n} X_i^{\gamma_i} = P_n \tag{18}$$

$$0 \le X_i \le \infty$$

where P_n and the γ_i are arbitrary constants. Although this constraint does not appear to conform to the type of constraint in eqn. 1, it can be readily cast into this form:

$$\int dV_n(P_n) = \int_0^{\infty} \prod_{i=1}^{n} dX_i \; \delta\left[P_n - \exp\left\{\sum_{i=1}^{n} \gamma_i \log X_i\right\}\right] . \tag{19}$$

Using k=1 in eqn. 3, we have

$$\int dV_n(P_n) = \int dV_1(P_1)\, dV_{n-1}(P_{n-1})\, \delta(P_n - P_k P_{n-k})\, dP_k\, dP_{n-k} \tag{20}$$

and

$$\int dV_1(P_1) = \int_0^{\infty} dX_1\, \delta(P_1 - X_1^{\gamma_1}) = \frac{1}{\gamma_1} P_1^{\frac{1-\gamma_1}{\gamma_1}} \tag{21}$$

with $X_1 = P_1^{\frac{1}{\gamma_1}}$, $\gamma_1 \neq 0$. For the trivial case $\gamma_1 = 0$, we take $X_1 = P_1$. In what follows we assume the γ_i are non-zero.

Substituting eqn. 21 into eqn. 20, and integrating over P_{n-1}

$$\int dV_n(P_n) = \frac{1}{\gamma_1}\int P_1^{\frac{1-2\gamma_1}{\gamma_1}}\, dP_1 dV_{n-1}\left(\frac{P_n}{P_1}\right) \tag{22}$$

Repeating this procedure n-1 times, one obtains

$$\int dV_n(P_n) = \frac{1}{\prod_{i=1}^{n} \gamma_i}\int_0^{\infty}\left[\prod_{i=1}^{n-1} dP_i P_i^{\frac{1-2\gamma_i}{\gamma_i}}\right]\left(\frac{P_n}{\prod_{i=1}^{n-1} P_i}\right)^{\frac{1-\gamma_n}{\gamma_n}}$$

with

$$X_i = P_i^{\frac{1}{\gamma_i}} \qquad 1 \le i \le n-1$$

$$X_n = \left(\frac{P_n}{\prod_{i=1}^{n-1} P_i} \right)^{\frac{1}{\gamma_n}}$$

Therefore

$$\int dV_n(P_n) = \frac{P_n^{\frac{1-\gamma_n}{\gamma_n}}}{n \prod_{i=1}^{n} \gamma_i} \prod_{i=1}^{n-1} \mu\,(\gamma_i,\gamma_n) \int_0^{\infty} d\eta_i \tag{23}$$

where

$$\mu(\gamma_i,\gamma_n) = \begin{cases} 1 & (\gamma_i = \gamma_n) \\ \dfrac{\gamma_i\gamma_n}{\gamma_n-\gamma_i} & (\gamma_i \neq \gamma_n) \end{cases}$$

and

$$X_i = \left\{ \begin{array}{ll} \log \eta_i & (\gamma_i = \gamma_n) \\ \eta_i^{\frac{\gamma_n}{\gamma_n-\gamma_i}} & (\gamma_i \neq \gamma_n) \end{array} \right\} \quad 1 \le i \le n-1 \tag{24}$$

$$X_n = P_n^{\frac{1}{\gamma_n}} \prod_{i=1}^{n-1} X_i^{-\frac{\gamma_i}{\gamma_n}} .$$

The infinite range of the independent variables can be compressed to the interval (0,1) using standard techniques.[6] It can be easily verified from eqn. 24 that the X_i satisfy the constraint (eqn. 18), and that the constant Jacobian for this volume preserving transformation is given by eqn. 23.

Multiple Equality Constraints

The technique discussed in the previous section can be extended to the problem of several simultaneous equality constraints. Consider a multiply constrained n-dimensional integration of the form

$$\int dV_n(C_{n,}^{(1)}\ldots, C_n^{(m)}) = \int \prod_{i=1}^{n} dX_i \prod_{j=1}^{m} \delta\left\{C_n^{(j)} - F_j\left[\sum_{i=1}^{n} f_i^{(j)}(X_i)\right]\right\} \quad (25)$$

where m is the number of simultaneous constraints and the F_j and $f_i^{(j)}$ are as defined for eqn. 1. This represents an integration over an n-m dimensional manifold embedded in the n-dimensional space. Proceeding as for the case of a single equality constraint, one obtains in analogy to eqn. 3

$$\begin{aligned} \int dV_n(C_{n,}^{(1)}\ldots, C_n^{(m)}) &= \int \prod_{j=1}^{m} dC_k^{(j)} dC_{n-k}^{(j)} \\ &\quad \times \delta\{C_n^j - F_j\,[F_j^{-1}(C_k^{(j)}) + F_j^{-1}(C_{n-k}^{(j)})]\} \\ &\quad \times \quad dV_k(C_{k,}^{(1)}\ldots C_k^{(m)})\, dV_{n-k}(C_{n-k,}^{(1)}\ldots C_{n-k}^{(m)}) \,. \end{aligned} \quad (26)$$

As before, eqn. 26 provides a framework for deriving recursive procedures that utilize known solutions in lower dimensional cases to solve higher dimensional problems. Here one has the same flexibility in the degree by which the dimensionality can be reduced at each step ($1 \le k \le n-1$), and the largest reduction that can be achieved is n-m two-dimensional constrained integrations. In this case, however, owing to the increased number of constraints, it is considerably more difficult to obtain the lower dimensional solutions. We illustrate the application of eqn. 26 with a single example.

C. Intersection of Sphere and Ellipse Transform (ISET)

Consider the integral over an n-2 dimensional manifold defined by the intersection of the surfaces of an n-dimensional sphere and ellipse,

$$dV_n(r_n, P_n) = \int \prod_{i=1}^{n} dX_i \;\; \delta\left(r_n - \sqrt{\sum_{i=1}^{n} X_i^2}\right) \delta\left(P_n - \sqrt{\sum_{i=1}^{n} (\alpha_i X_i)^2}\right) . \tag{27}$$

For n=2, one has the result

$$dV_2\,(r_2,\, P_2) \, . \quad = \frac{4P_2 r_2}{\sqrt{\left(\frac{P_2^2}{r_2^2} - \alpha_2^2\right)\left(\alpha_1^2 - \frac{P_2^2}{r_2^2}\right)}} \tag{28}$$

where the conditions

$$\alpha_2^2 < \frac{P_2^2}{r_2^2} < \alpha_1{}^2$$

are required for the sphere and ellipse to intersect. The transformation in this case becomes

$$\begin{aligned} X_1 &= \pm\, r_2 \sqrt{\left(\frac{P_2^2}{r_2^2} - \alpha_2^2\right) \Big/ \left(\alpha_1^2 - \alpha_2^2\right)} \\ X_2 &= \pm\, r_2 \sqrt{\left(\alpha_1^2 - \frac{P_2^2}{r_2^2}\right) \Big/ \left(\alpha_1^2 - \alpha_2^2\right)} \end{aligned} \tag{29}$$

Then, using k=2 in eqn. 26, one obtains

$$\int dV_n(r_n,P_n) = 4 \int dV_{n-2}(r_{n-2},P_{n-2}) \frac{r_2\ P_2}{\sqrt{\left(\frac{P_2^2}{r_2^2} - \alpha_2^2\right)\left(\alpha_1^2 - \frac{P_2^2}{r_2^2}\right)}}$$

$$\times\ \delta\left(r_n - \sqrt{r_2^2 + r_{n-2}^2}\right)\ \delta\left(P_n - \sqrt{P_2^2 + P_{n-2}^2}\right)$$

$$\times\ dr_2 dr_{n-2}\ dP_2\ dP_{n-2} \quad .$$

Transforming to polar coordinates

$$r_2 = \zeta_1 \cos\ \theta_1 \qquad\qquad P_2 = \zeta_2 \cos\ \theta_2$$

$$r_{n-2} = \zeta_1 \sin\ \theta_1 \qquad\qquad P_{n-2} = \zeta_2 \sin\ \theta_2$$

and integrating over ζ_1 and ζ_2, one obtains

$$\int dV_n(r_n,P_n) = 16 r_n^2 P_n^2 \int_0^1 \frac{dV_{n-2}(r_n\mu_1,\ P_n\mu_2)\ d\mu_1 d\mu_2}{\sqrt{\left[\frac{P_n^2(1-\mu_2^2)}{r_n^2(1-\mu_1^2)} - \alpha_2^2\right]\left[\alpha_1^2 - \frac{P_n^2(1-\mu_2^2)}{r_n^2(1-\mu_1^2)}\right]}} \tag{30}$$

This represents a recursion relation for eliminating the constraints two dimensions at a time until, ultimately, one is left with an n-2 dimensional unconstrained integration in $\mu_1 \ldots \mu_{n-2}$. The resulting transformation from the μ's to the x's are analogous to eqn. 29.

Note that, unlike the single constraint case, the n-2 μ's are not independent but are pairwise correlated. In general, with m constraints the integration variables are m-wise correlated.

REFERENCES

1. J. Kowalik and M.R. Osborne, "Methods for Unconstrained Optimization Problems", American Elsevier (1968), Chapter 1.

2. Ibid, Chapter 5.

3. C.J. Everett and E.D. Cashwell, "A Monte Carlo Sampler", Los Alamos Report LA-5061-MS, October 1972.

4. K.S. Miller, "Multidimensional Gaussian Distributions", Wiley (1964), Chapter 1.

5. D. Shelupsky, Am. Math. Monthly, 69, No. 7, p. 644-646, (1962).

6. M.J. Box, Computer Journal 9, p. 67-77, (1966).