Chapter 3

Equality-Constrained Optimization: Normality Conditions

3.1 Introduction

This chapter deals with the simplest class of constrained-optimization problems, those subject to equality constraints. The problem statement at hand is

$$f(\mathbf{x}) \quad \rightarrow \quad \min_{\mathbf{x}} \tag{3.1a}$$

subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}_l \tag{3.1b}$$

where $\mathbf{h}$ is a *smooth* l-dimensional vector function[1] of the n-dimensional vector argument $\mathbf{x}$, $\mathbf{0}_l$ denoting the l-dimensional zero vector. Moreover, $l < n$, for an n-dimensional design vector $\mathbf{x}$.

In the sequel, we will resort to concepts of linear transformations of vector spaces, such as *rotations*, *reflections* and *projections*. We start with an outline of these concepts.

The main outcome is the derivation of the *normality conditions* of the problems at hand. We derive these in two forms: the *direct form*, in terms of the gradients of the objective function $f(\mathbf{x})$ to be minimized and of the constraints (3.1b), and the *dual form*, in terms of an *orthogonal complement* of the gradient of $\mathbf{h}$ w.r.t. $\mathbf{x}$.

[1]Smoothness implies that $\mathbf{h}(\mathbf{x})$ is continuous and has a continuous gradient w.r.t. $\mathbf{x}$.

As a special case, that lends itself to a closed-form solution, we study *minimum-norm* problems, whereby a *weighted* Euclidean norm of the design vector is to be minimized subject to l linear equality constraints. In this vein, we introduce the *right Moore-Penrose generalized inverse*.

3.2 Background on Linear Transformations

The general form of a *linear transformation* mapping a vector space $\mathcal{U}$ of dimension n into an m-dimensional vector space $\mathcal{V}$ is

$$\mathbf{v} = \mathbf{L}\mathbf{u} \tag{3.2}$$

where $\mathbf{u}$ and $\mathbf{v}$ are n- and m-dimensional vectors, respectively, with $\mathbf{u} \in \mathcal{U}$ and $\mathbf{v} \in \mathcal{V}$. Apparently, $\mathbf{L}$ is a $m \times n$ matrix.

We distinguish two vector subspaces associated with $\mathbf{L}$, namely,

The *range* of $\mathbf{L}$, denoted by $\mathcal{R}(\mathbf{L})$: the set of vectors $\mathbf{v}$ that are *images* of $\mathbf{L}$ under transformation (3.2). Notice that, if the n columns of $\mathbf{L}$ are not linearly independent, then $\mathcal{R}(\mathbf{L})$ is not all of $\mathcal{V}$, but only a *proper* subspace of it, of dimension $m' < n$, i.e., $\mathcal{R}(\mathbf{L}) \subset \mathcal{V}$. The dimension of $\mathcal{R}(\mathbf{L})$ is denoted by $\rho(\mathbf{L})$.

The *nullspace* or *kernel* of $\mathbf{L}$, denoted by $\mathcal{N}(\mathbf{L})$: the set of all nonzero vectors $\mathbf{u}$ of $\mathcal{U}$ that are mapped by $\mathbf{L}$ into $\mathbf{0}_m$, the zero of $\mathcal{V}$. The dimension of $\mathcal{N}$ is termed the *nullity* of $\mathbf{L}$, and is denoted by $\nu(\mathbf{L})$. Obviously, $\nu < n$, with $\nu = n$ occurring only when $\mathbf{L} = \mathbf{O}_{mn}$, $\mathbf{O}_{mn}$ denoting the $m \times n$ zero matrix.

A fundamental result of linear algebra follows:

$$\rho(\mathbf{L}) + \nu(\mathbf{L}) = n \tag{3.3}$$

The most frequent linear transformations used in optimum design are studied in the balance of this section. They all pertain to square matrices.

3.2.1 Rotations

A rotation $\mathbf{Q}$ is an *orthogonal* transformation of $\mathcal{U}$ into itself whose determinant is positive. Orthogonality requires that the inverse of $\mathbf{Q}$ be its transpose, i.e.,

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{1} \tag{3.4}$$

where $\mathbf{1}$ denotes the $n \times n$ identity matrix. Hence, taking the determinant of both sides of the above equation,

$$\det(\mathbf{Q}\mathbf{Q}^T) = \det(\mathbf{Q}^T\mathbf{Q}) = \det(\mathbf{Q})^T\det(\mathbf{Q}) = [\det(\mathbf{Q})]^2 = 1$$

whence

$$\det(\mathbf{Q}) = \pm 1$$

A *proper* orthogonal matrix $\mathbf{Q}$ is one whose determinant is positive, and hence,

$$\det(\mathbf{Q}) = +1 \tag{3.5}$$

Proper orthogonal transformations of $\mathcal{U}$ into itself represent rotations about the origin of $\mathcal{U}$.

The best-known rotations are those in two and three dimensions. Thus, for two dimensions, the 2×2 matrix $\mathbf{Q}$ rotating vectors through an angle ϕ ccw takes the form

$$\mathbf{Q} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \tag{3.6a}$$

which can be expressed alternatively as

$$\mathbf{Q} = (\cos\phi)\mathbf{1} + (\sin\phi)\mathbf{E} \tag{3.6b}$$

with $\mathbf{1}$ defined as the 2×2 identity matrix and $\mathbf{E}$ as a skew-symmetric matrix, namely,

$$\mathbf{E} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{3.6c}$$

In three dimensions, the rotation matrix takes the form

$$\mathbf{Q} = \mathbf{e}\mathbf{e}^T + \cos\phi(\mathbf{1} - \mathbf{e}\mathbf{e}^T) + \sin\phi\mathbf{E} \tag{3.7a}$$

where $\mathbf{e}$ is the unit vector indicating the direction of the axis about which the rotation takes place. Moreover, $\mathbf{E}$ is now defined as the *cross-product matrix* (CPM) of vector $\mathbf{e}$, expressed as

$$\mathbf{E} \equiv \mathrm{CPM}(\mathbf{e}) \tag{3.7b}$$

and should not be confused with its two-dimensional counterpart, which is also skew-symmetric.

The cross-product matrix is defined only for three-dimensional vectors. Let $\mathbf{a}$ and $\mathbf{v}$ be two arbitrary three-dimensional vectors. We define

$$\mathrm{CPM}(\mathbf{a}) \equiv \frac{\partial(\mathbf{a} \times \mathbf{v})}{\partial \mathbf{v}} \equiv \mathbf{A} \tag{3.8}$$

Because of the relation $\mathbf{a} \times \mathbf{v} = -\mathbf{v} \times \mathbf{a}$, the CPM is skew-symmetric.

Rotations in higher dimensions can be defined as well, but then, the axis and the angle of rotation are not unique.

3.2.2 Reflections

An instance of reflections, namely, $n \times n$ Householder reflections $\mathbf{H}_i$, was introduced in terms of an n-dimensional vector $\mathbf{u}_i$ in eq.(2.49). The geometric interpretation of this transformation is straightforward: $\mathbf{H}_i$ reflects vectors in n-dimensional space about a plane Π_i of unit normal $\mathbf{u}_i/\|\mathbf{u}_i\|$ passing through the origin; all vectors normal to $\mathbf{u}_i$ remain *immutable* under $\mathbf{H}_i$.

In two dimensions, a reflection $\mathbf{R}$ about a line passing through the origin normal to the unit vector $\mathbf{e}$ maps vectors $\mathbf{p}$ into $\mathbf{p}'$ in the form

$$\begin{aligned}\mathbf{p}' &= \mathbf{p} - 2(\mathbf{p}^T\mathbf{e})\mathbf{e} \\ &= (\mathbf{1} - 2\mathbf{e}\mathbf{e}^T)\mathbf{p} \equiv \mathbf{R}\mathbf{p}\end{aligned}$$

and hence, the reflection $\mathbf{R}$ sought is given by

$$\mathbf{R} = \mathbf{1} - 2\mathbf{e}\mathbf{e}^T \tag{3.9}$$

In three dimensions, the reflection about a plane passing through the origin, of unit normal $\mathbf{e}$, takes exactly the same form as $\mathbf{R}$ in the two-dimensional case, eq.(3.9). As a matter of fact, $\mathbf{H}_i$ of eq.(2.49) takes the same form as $\mathbf{R}$ of eq.(3.9), if we replace $\mathbf{u}_i/\|\mathbf{u}_i\|$ by $\mathbf{e}$.

In all foregoing instances, the reflections are represented by symmetric matrices, and are hence termed *pure reflections*. However, this need not always be the case, for reflections can combine with rotations, thereby yielding a new reflection—notice that the product of a rotation by a pure reflection is a reflection!—but this time, the matrix representing the reflection is no longer symmetric. A rotation can be distinguished from a reflection by the sign of its determinant. Indeed, for any $n \times n$ reflection $\mathbf{R}$, we have

$$\det(\mathbf{R}) = -1 \tag{3.10}$$

3.2.3 Projections

Henceforth, a *projection* $\mathbf{P}$ means an *orthogonal projection* onto a plane in n dimensions, which we call the *projection plane*. When $n = 2$, the "projection plane" becomes a line in the plane.

Let us consider a plane Π in an n-dimensional space, of unit normal $\mathbf{n}$. Any point P in this space is given by its n-dimensional position vector $\mathbf{p}$. Let the projection of P onto Π be P', which is given by its position vector $\mathbf{p}'$, namely,

$$\mathbf{p}' = \mathbf{p} - (\mathbf{n}^T\mathbf{p})\mathbf{n} = (\mathbf{1} - \mathbf{n}\mathbf{n}^T)\mathbf{p} \equiv \mathbf{P}\mathbf{p} \tag{3.11}$$

where $\mathbf{P}$ was obviously defined as

$$\mathbf{P} \equiv \mathbf{1} - \mathbf{n}\mathbf{n}^T \tag{3.12}$$

Matrix $\mathbf{P}$ is also called a *projector*. A projector $\mathbf{P}$ is represented by a symmetric, singular, *idempotent* matrix. Symmetry is obvious; singularity is less so, but rather straightforward. To prove that $\mathbf{P}$ is singular, all we have to do is prove that its nullspace is non-empty. However, this is so because all vectors $\mathbf{r}$ of the form $\alpha\mathbf{n}$, for a scalar α, are mapped by $\mathbf{P}$ onto the zero vector. Indeed,

$$\mathbf{P}\mathbf{r} = \alpha\mathbf{P}\mathbf{n} = \alpha(\mathbf{1} - \mathbf{n}\mathbf{n}^T)\mathbf{n} = \alpha(\mathbf{n} - \mathbf{n}) = \mathbf{0}$$

A matrix is idempotent of degree k when it equals its kth power, but is different from any lower power. When $k = 1$, the degree is self-understood. To prove idempotency, let is calculate

$$\begin{aligned}\mathbf{P}^2 &= (\mathbf{1} - \mathbf{n}\mathbf{n}^T)(\mathbf{1} - \mathbf{n}\mathbf{n}^T) = \mathbf{1} - 2\mathbf{n}\mathbf{n}^T + \mathbf{n}\underbrace{\mathbf{n}^T\mathbf{n}}_{=1}\mathbf{n}^T \\ &= \mathbf{1} - \mathbf{n}\mathbf{n}^T \equiv \mathbf{P}\end{aligned}$$

thereby completing the proof.

The foregoing projection has a nullity of 1, its nullspace being spanned by vector $\mathbf{n}$. In three-dimensional space, we can have projections onto a subspace of dimension 1, namely, a line $\mathcal{L}$ passing through the origin and parallel to the unit vector $\mathbf{e}$. In this case, the projection P' of P onto $\mathcal{L}$ is given by

$$\mathbf{p}' = (\mathbf{p}^T\mathbf{e})\mathbf{e} \equiv \mathbf{e}(\mathbf{e}^T\mathbf{p}) = (\mathbf{e}\mathbf{e}^T)\mathbf{p}$$

whence the projection $\mathbf{P}$ sought takes the form:

$$\mathbf{P} = \mathbf{e}\mathbf{e}^T \tag{3.13}$$

Notice that this projection is symmetric, singular and idempotent as well, its nullspace being of dimension two. Indeed, we can find two mutually-orthogonal unit vectors

$\mathbf{f}$ and $\mathbf{g}$, lying in a plane normal to $\mathbf{e}$, which are mapped by $\mathbf{P}$ onto the zero vector. These two linearly-independent vectors lie in the nullspace of $\mathbf{P}$. For n dimensions, the projection "plane" can in fact be a subspace of dimension $\nu \leq n-1$.

Also notice that the projection of eq.(3.11) maps vectors in three-dimensional space onto the nullspace of the *rank-one matrix* $\mathbf{n}\mathbf{n}^T$, while that of eq.(3.13) does so onto the range of the rank-one matrix $\mathbf{e}\mathbf{e}^T$. Now, the range of this matrix is the nullspace of a matrix $\mathbf{A}$ defined as

$$\mathbf{A} \equiv \begin{bmatrix} \mathbf{f}^T \\ \mathbf{g}^T \end{bmatrix} \tag{3.14}$$

where $\mathbf{f}$ and $\mathbf{g}$ are mutually orthogonal unit vectors normal to $\mathbf{e}$. Then, we can define a projector $\mathbf{P}$ in the form

$$\mathbf{P} = \mathbf{1} - \mathbf{A}^T\mathbf{A} = \mathbf{1} - (\mathbf{f}\mathbf{f}^T + \mathbf{g}\mathbf{g}^T) \tag{3.15}$$

This projector maps three-dimensional vectors onto the nullspace of $\mathbf{A}$, which is vector $\mathbf{e}$, as the reader can readily verify.

In general, if we have a full-rank $m \times n$ matrix $\mathbf{A}$, with $m < n$, then, $\text{rank}(\mathbf{A}) = \min\{m, n\} = m$. This means that the m n-dimensional rows of $\mathbf{A}$ are linearly independent. By virtue of the basic relation (3.3), then, $\nu = n - m$. A projector that maps n-dimensional vectors onto the nullspace of $\mathbf{A}$ is defined below:

$$\mathbf{P} = \mathbf{1} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A} \tag{3.16}$$

Note that, by virtue of the definition of $\mathbf{f}$ and $\mathbf{g}$, matrix $\mathbf{A}$ of eq.(3.14) produces $\mathbf{A}\mathbf{A}^T = \mathbf{1}$.

Exercise 3.2.1 *Prove that* $\mathbf{P}$*, as given by eq.(3.16), is a projection; then prove that this projection maps n-dimensional vectors onto the nullspace of* $\mathbf{A}$*.*

Example 3.2.1 *Let*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \equiv \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix}$$

The nullspace of $\mathbf{A}$ *is spanned by a unit vector* $\mathbf{u}$ *that can be found as*

$$\mathbf{u} \equiv \frac{\mathbf{b}}{\|\mathbf{b}\|}, \quad \mathbf{b} \equiv \mathbf{a}_1^T \times \mathbf{a}_2^T$$

The projector $\mathbf{P}$ *mapping vectors in three-dimensional space onto the nullspace of* $\mathbf{A}$*, spanned by* $\mathbf{u}$*, is given by*

$$\mathbf{P} = \mathbf{1} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A} = \frac{1}{3}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

In this case,

$$\mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u} = \frac{\sqrt{3}}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Notice that the image of any vector $\mathbf{p} = [\, x,\, y,\, z\,]^T$ *under* $\mathbf{P}$ *can be expressed as the product of a scalar times* $\mathbf{u}$*:*

$$\begin{aligned} \mathbf{Pp} &= \frac{1}{3} \begin{bmatrix} x - y - z \\ -x + y + z \\ -x + y + z \end{bmatrix} = \frac{1}{3}(-x + y + z) \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \\ &= \frac{\sqrt{3}}{3}(-x + y + z)\mathbf{u} \end{aligned}$$

3.3 Equality-Constrained Minimization

3.3.1 First-Order Normality Conditions

We now derive the normality conditions of problem (3.1a). To this end, we resort to *Lagrange multipliers* $\lambda_1, \lambda_2, \ldots, \lambda_l$, one for each scalar constraint $h_i(\mathbf{x}) = 0$, and group them in the l-dimensional array $\boldsymbol{\lambda}$. Upon adjoining the l constraints to the objective function $f(\mathbf{x})$, we obtain the *Lagrangian* $F(\mathbf{x};\, \boldsymbol{\lambda})$ that we aim at minimizing under no constraints, while choosing $\boldsymbol{\lambda}$ in such a way that the l equality constraints are satisfied. That is,

$$F(\mathbf{x};\, \boldsymbol{\lambda}) \equiv f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{h} \quad \rightarrow \quad \min_{\mathbf{x}, \boldsymbol{\lambda}} \tag{3.17}$$

subject to no constraints. We thus have transformed the equality-constrained minimization problem into an unconstrained one. We derive now the normality conditions of the problem at hand by recalling those of Chapter 2, requiring that a) the gradient of the objective function w.r.t. the design variables vanish and b) the Hessian of the objective function w.r.t. the same variables be positive-definite. However, note that we now have l additional variables besides the original n design variables. We thus have to augment the design-variable vector correspondingly, which we do by defining an augmented $(n + l)$-dimensional design vector $\mathbf{y}$:

$$\mathbf{y} \equiv \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} \tag{3.18}$$

Therefore, the unconstrained minimization problem (3.17) can be formulated in a more compact form, namely,

$$F(\mathbf{y}) \quad \rightarrow \quad \min_{\mathbf{y}} \tag{3.19}$$

subject to no constraints. The first-order normality conditions of the above problem are, thus,

$$\frac{\partial F}{\partial \mathbf{y}} = \begin{bmatrix} \partial F/\partial \mathbf{x} \\ \partial F/\partial \boldsymbol{\lambda} \end{bmatrix} = \mathbf{0}_{n+l} \tag{3.20a}$$

where $\mathbf{0}_{n+l}$ denotes the $(n+l)$-dimensional zero vector. Upon expansion, the above equation yields

$$\frac{\partial F}{\partial \mathbf{x}} = \mathbf{0}_n \tag{3.20b}$$

$$\frac{\partial F}{\partial \boldsymbol{\lambda}} = \mathbf{0}_l \tag{3.20c}$$

To gain insight into the geometric significance of the foregoing normality conditions, we expand the left-hand side of eq.(3.20b) componentwise:

$$\begin{aligned}
\frac{\partial F}{\partial x_1} &\equiv \frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial h_1}{\partial x_1} + \lambda_2 \frac{\partial h_2}{\partial x_1} + \cdots + \lambda_l \frac{\partial h_l}{\partial x_1} = 0 \\
\frac{\partial F}{\partial x_2} &\equiv \frac{\partial f}{\partial x_2} + \lambda_1 \frac{\partial h_1}{\partial x_2} + \lambda_2 \frac{\partial h_2}{\partial x_2} + \cdots + \lambda_l \frac{\partial h_l}{\partial x_2} = 0 \\
&\vdots \\
\frac{\partial F}{\partial x_n} &\equiv \underbrace{\frac{\partial f}{\partial x_n}}_{\nabla f} + \lambda_1 \frac{\partial h_1}{\partial x_n} + \lambda_2 \frac{\partial h_2}{\partial x_n} + \cdots + \lambda_l \frac{\partial h_l}{\partial x_n} = 0
\end{aligned} \tag{3.21}$$

where the the first term of the ith equation can be readily identified as the ith component of $\nabla f = \partial f/\partial \mathbf{x}$. The remaining terms of the same equation can be identified as the ith component of an inner product p_i defined as

$$p_i \equiv [\,\partial h_1/\partial x_i \quad \partial h_2/\partial x_i \quad \cdots \quad \partial h_l/\partial x_i\,] \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_l \end{bmatrix}$$

Therefore,

$$\mathbf{p} = \underbrace{\begin{bmatrix} \partial h_1/\partial x_1 & \partial h_2/\partial x_1 & \cdots & \partial h_l/\partial x_1 \\ \partial h_1/\partial x_2 & \partial h_2/\partial x_2 & \cdots & \partial h_l/\partial x_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial h_1/\partial x_n & \partial h_2/\partial x_n & \cdots & \partial h_l/\partial x_n \end{bmatrix}}_{(\nabla \mathbf{h})^T:\; n \times l} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_l \end{bmatrix}}_{\boldsymbol{\lambda}} \tag{3.22}$$

which can be readily identified, in turn, as the product $(\nabla\mathbf{h})^T\boldsymbol{\lambda}$. The first n normality conditions displayed in eq.(3.20b) thus amount to

$$\nabla f + \mathbf{J}^T\boldsymbol{\lambda} = \mathbf{0}_n \tag{3.23a}$$

where $\mathbf{J} \equiv \nabla\mathbf{h}$ is the gradient of $\mathbf{h}$ with respect to $\mathbf{x}$. The remaining l normality conditions, displayed in eq.(3.20c), yield nothing but the constraints themselves, namely

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}_l \tag{3.23b}$$

Equation (3.23a) is the vector representation of the *first-order normality conditions* (FONC) sought.

What eq.(3.23a) represents has a geometric significance that will be made apparent upon rewriting it in the alternative form

$$\mathbf{J}^T\boldsymbol{\lambda} = -\nabla f \tag{3.24}$$

The foregoing equation states that, at a stationary point $\mathbf{x}_0$, $-\nabla f$, or ∇f for that matter, lies in the range of the transpose of the gradient of the constraints. Notice that the range $\mathcal{J}'$ of $\mathbf{J}^T$ is a subspace of the n-dimensional space of design variables. In fact, $\dim[\mathcal{R}(\mathcal{J}')] = l < n$, for this subspace is spanned by l linearly-independent vectors, the columns of $\mathbf{J}^T$, or the n-dimensional rows of $\mathbf{J}$.

Algebraically, what eq.(3.24) represents is an overdetermined system of n linear equations in the $l < n$ unknowns $\{\lambda_i\}_1^l$. The normality condition then states that the least-square approximation of this overdetermined system yields a zero error. That is, at a stationary point, the n $(> l)$ equations (3.24) become all consistent. Note that the least-square approximation $\boldsymbol{\lambda}_0$ of the foregoing equations can be expressed in terms of the left Moore-Penrose generalized inverse of $\mathbf{J}^T$, namely,

$$\boldsymbol{\lambda}_0 = -(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla f \tag{3.25}$$

The least-square error $\mathbf{e}_0$ of this approximation is thus

$$\begin{aligned}\mathbf{e}_0 &= \mathbf{J}^T\boldsymbol{\lambda}_0 - (-\nabla f) = -\mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}\nabla f + \nabla f \\ &= [\mathbf{1} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}]\nabla f\end{aligned} \tag{3.26}$$

with $\mathbf{1}$ denoting the $n \times n$ identity matrix.

We can now express the first-order normality condition (3.24) in yet one more alternative form:

$$[\mathbf{1} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}]\nabla f = \mathbf{0}_n \tag{3.27}$$

The matrix inside the brackets in the foregoing equation can be readily identified as a projector. This projector maps vectors in $\mathbb{R}^n$ onto the nullspace of $\mathbf{J}$. In other words, *at a stationary point* P_0 *the gradient of the objective function need not vanish; only its projection onto the nullspace of the gradient of the constraints must vanish*, which is an alternative form of stating the first-order normality condition. Sometimes the product $\overline{\nabla} f$, defined as

$$\nabla f \equiv [\mathbf{1} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}]\nabla f \in \mathbb{R}^n \tag{3.28}$$

is referred to as the *constrained gradient*. The FONC (3.27) can then be simply stated as: *At a stationary of the equality-constrained problem (3.1a & b), the constrained gradient vanishes.*

Exercise 3.3.1

Prove that

$$\mathbf{P} \equiv \mathbf{1} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J} \tag{3.29}$$

indeed maps vectors in $\mathbb{R}^n$ onto the nullspace of $\mathbf{J}$ and that $\mathbf{P}$ is a projector.

Dual Form of the FONC

One more form of the FONC of the problem at hand is now derived in what we can term *dual form*. To this end, we realize that the solution sought $\mathbf{x}_0$ must lie in a subset of the space $\mathbb{R}^n$ of design variables, of reduced dimension $n - l$, which contains all design vectors $\mathbf{x}$ satisfying the constraints. Now, this set need not be a vector space, and in general it is not. Indeed, if the constraints are nonlinear, then the sum of two distinct vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ will most likely not satisfy the constraints, even if these two designs do so independently. Neither does the zero vector satisfy the constraints, which thus disqualify the set from being a subspace of $\mathbb{R}^n$. What we have as a *feasible subset* of the design space is a *manifold* $\mathcal{F}$, i.e., a smooth surface embedded in $\mathbb{R}^n$. We shall term this subset the *feasible manifold*.

Finding $\mathcal{F}$ may be a tremendous task when the constraints are nonlinear and algebraically complicated. The good news is that we do not actually need the feasible manifold to obtain a feasible solution. What we really need is a *feasible subspace* tangent to the said manifold at a feasible point. We discuss below how to obtain this subspace. Assume that we have a feasible point P_F, of position vector $\mathbf{x}_F$, i.e.,

$$\mathbf{h}(\mathbf{x}_F) = \mathbf{0}_l \tag{3.30}$$

An arbitrary "move" $\Delta\mathbf{x}$ from $\mathbf{x}_F$ will most likely take P_F away from the constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}_l$. What we would like to find is a *feasible move*, i.e., a vector $\Delta\mathbf{x}_F$ such that $\mathbf{x}_F + \Delta\mathbf{x}_F$ still verifies the constraints, i.e.,

$$\mathbf{h}(\mathbf{x}_F + \Delta\mathbf{x}_F) = \mathbf{h}(\mathbf{x}_F) + \mathbf{J}(\mathbf{x}_F)\Delta\mathbf{x}_F + \text{HOT} = \mathbf{0}_l \tag{3.31}$$

where HOT stands for "higher-order terms". Since we assumed at the outset that $\mathbf{x}_F$ is feasible, eq.(3.30), we have, from the foregoing equation and to a first-order approximation, i.e., neglecting HOT,

$$\mathbf{J}(\mathbf{x}_F)\Delta\mathbf{x}_F = \mathbf{0}_l \tag{3.32}$$

Moreover, $\mathbf{J}$ is of $l \times n$, with $l < n$, and hence, it is possible to find $n - l$ linearly independent vectors of $\mathbb{R}^n$ lying in $\mathcal{N}(\mathbf{J}(\mathbf{x}_F))$, i.e., in the nullspace of $\mathbf{J}(\mathbf{x}_F)$. These vectors can be produced in many ways. We will not dwell, for the time being, in the pertinent computing methods, but rather on the concepts behind the production of those $n - l$ vectors. Let us thus assume that we have found such $n - l$ linearly independent vectors, arrayed in the $n \times n'$ matrix $\mathbf{L}$, with $n' \equiv n - l$, and hence,

$$\mathbf{J}\mathbf{L} = \mathbf{O}_{ln'} \tag{3.33}$$

matrix $\mathbf{L}$ being termed an *orthogonal complement* of $\mathbf{J}$ and $\mathbf{O}_{ln'}$ denoting the $l \times n'$ zero matrix.

Now, if we define

$$\Delta\mathbf{x}_F = \mathbf{L}\Delta\mathbf{u} \tag{3.34a}$$

for arbitrary $\Delta\mathbf{u} \in \mathbb{R}^{n'}$, we will have

$$\mathbf{J}\Delta\mathbf{x}_F = \mathbf{J}\mathbf{L}\Delta\mathbf{u} = \mathbf{0}_{n'} \tag{3.34b}$$

the "move" $\Delta\mathbf{x}_F$ thus verifying the constraints to a first degree. Now, the first-order normality condition of the problem at hand can be cast in the form

$$\Delta f \equiv (\nabla f)^T \Delta\mathbf{x}_F = (\nabla f)^T \mathbf{L}\Delta\mathbf{u} = (\mathbf{L}^T \nabla f)^T \Delta\mathbf{u} = 0 \quad \forall \quad \Delta\mathbf{u}$$

Hence, the alternative form of the FONC is

$$\mathbf{L}^T \nabla f = \mathbf{0}_{n'} \tag{3.35}$$

That is, at a stationary point, the gradient of f need not vanish; it must lie in the nullspace of $\mathbf{L}^T$, i.e., in the range of $\mathbf{L}$. We can thus call $\mathbf{L}^T\nabla f$ the *feasible gradient*, and represent it by $\nabla_u f$, i.e.,

$$\nabla_u f = \mathbf{L}^T \nabla f \tag{3.36}$$

which is a $(n-l)$-dimensional vector. Notice that, from eq.(3.34a), $\mathbf{L}$ has the differential interpretation

$$\mathbf{L} = \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \tag{3.37}$$

and hence, the FONC (3.35) can be restated as

$$\left(\frac{\partial \mathbf{x}}{\partial \mathbf{u}}\right)^T \left(\frac{\partial f}{\partial \mathbf{x}}\right) \equiv \frac{\partial f}{\partial \mathbf{u}} = \mathbf{0}_{n'} \quad \text{or} \quad \nabla_u f = \mathbf{0}_{n'} \tag{3.38}$$

That is, the FONC (3.38) states that, at a stationary point of problem (3.1a & b), the gradient of $f(\mathbf{x})$ w.r.t. the vector of independent design variables $\mathbf{u}$ vanishes.
Remark: When comparing the two forms of the FONC, eqs.(3.27) and (3.35), the simplicity of the latter with respect to the former is apparent. This simplicity, however, is more than formal, for eq.(3.27) involves n scalar equations, while eq.(3.35) involves only $n-l$ scalar equations.

3.3.2 Second-Order Normality Conditions

The second-order normality conditions of the problem at hand require that the Hessian of the Lagrangian $F(\mathbf{x};\boldsymbol{\lambda})$ with respect to the $(n+l)$-dimensional *augmented* design-variable vector $\boldsymbol{\xi} \equiv [\mathbf{x}^T \ \boldsymbol{\lambda}^T]^T$ be positive-definite. The said Hessian takes the form

$$\frac{\partial^2 F}{\partial \boldsymbol{\xi}} = \begin{bmatrix} \nabla\nabla f + \partial(\mathbf{J}^T\boldsymbol{\lambda})/\partial\mathbf{x} & \mathbf{J}^T \\ \mathbf{J} & \mathbf{O}_l \end{bmatrix} \tag{3.39}$$

where, as usual, $\nabla\nabla f$ represents the Hessian of $f(\mathbf{x})$ with respect to $\mathbf{x}$, and $\mathbf{O}_l$ denotes the $l \times l$ zero matrix. Notice that computing the above Hessian requires the computation of $\nabla\nabla f$ and $\partial(\mathbf{J}^T\boldsymbol{\lambda})/\partial\mathbf{x}$, which involves the second derivatives of $\mathbf{h}(\mathbf{x})$ with respect to $\mathbf{x}$. Moreover, in order to verify the second-order normality conditions, the $n+l$ eigenvalues of $\partial^2 F/\partial\boldsymbol{\xi}$ must be computed. Thus, the second-order normality conditions in this form are extremely costly to verify.

Alternatively, we resort to the form (3.34a) of $\Delta\mathbf{x}_F$ and assume that we have found a stationary point verifying eq.(3.35). Upon expansion of $f(\mathbf{x}_F + \Delta\mathbf{x}_F)$ to second order, we obtain

$$\Delta f = f(\mathbf{x}_F + \Delta\mathbf{x}_F) - f(\mathbf{x}_F) \approx (\nabla f|_{x=\mathbf{x}_0})^T \Delta\mathbf{x}_F + \frac{1}{2}(\Delta\mathbf{x}_F)^T \nabla\nabla f\Big|_{x=\mathbf{x}_0} \Delta\mathbf{x}_F$$

which must be positive if the current value $\mathbf{x}_0$ is a minimum. Thus, upon substituting eq.(3.34a) into the above expression, we obtain

$$\Delta f \approx (\nabla f|_{x=\mathbf{x}_0})^T \mathbf{L}\Delta\mathbf{u} + \frac{1}{2}\Delta\mathbf{u}^T \mathbf{L}^T \nabla\nabla f\Big|_{x=\mathbf{x}_0} \mathbf{L}\Delta\mathbf{u} > 0$$

Now, since we have assumed that the FONC holds at the stationary point $\mathbf{x}_0$, eq.(3.35) holds, and hence,

$$(\nabla f|_{x=\mathbf{x}_0})^T \mathbf{L}\Delta\mathbf{u} = 0$$

the *second-order normality condition* (SONC) thus becoming

$$\Delta\mathbf{u}^T \mathbf{L}^T \nabla\nabla f\big|_{x=\mathbf{x}_0} \mathbf{L}\Delta\mathbf{u} > 0 \quad \forall \quad \Delta\mathbf{u} \tag{3.40}$$

We term the product $\mathbf{L}^T \nabla\nabla f\big|_{x=\mathbf{x}_0} \mathbf{L}$ the *feasible Hessian* of f. That is, *a stationary point* $\mathbf{x}_0$ *is a local minimum if its feasible Hessian is positive-definite. As a consequence, then, at a minimum, the Hessian itself need not be positive-definite, but its feasible component must be.*

We can now represent the $(n-l) \times (n-l)$ feasible Hessian as

$$\nabla_u \nabla_u f \equiv \mathbf{L}^T \nabla\nabla f \mathbf{L} = \left(\frac{\partial \mathbf{x}}{\partial \mathbf{u}}\right)^T (\nabla\nabla f) \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \tag{3.41}$$

the SONC thus lending themselves to a more straightforward interpretation:

A stationary point $\mathbf{x}_0$ *of problem (3.1a & b) is a minimum iff the Hessian of* $f(\mathbf{x})$ *w.r.t. the independent design-variable vector* $\mathbf{u}$ *is positive-definite.*

Remark: At a minimum, the Hessian of f w.r.t. $\mathbf{x}$ need not be positive-definite. However, if $\nabla\nabla f$ is positive-definite, then $\nabla_u \nabla_u f$ is necessarily positive-definite as well.

Example 3.3.1 (The Design of a Manipulator for a Given Reach)

In designing the manipulator of Fig. 3.1 (Angeles, 2002), we want to find the value of the length a that will produce the reach of a Puma 560 robot, namely, 0.8772 m.

It is apparent that the maximum reach is independent of θ_1*, the angle of rotation of the first joint, for motions about the first joint do not affect the reach. So, we lock the first joint and, in the posture of Fig. 3.1, rotate the third joint through one full turn, point C thus describing a circle* $\mathcal{C}$ *of radius a lying in the* Y_1*-*Z_1 *plane, with centre at point* O_3' *of coordinates* $(0,\ a,\ -a)$*. Next, upon performing a full rotation of the second joint, the circle describes a toroid of axis* Z_2*, the problem now reducing to one of finding the point of the surface of the toroid lying the farthest from the* Z_1 *axis. Figure 3.2 includes side views of circle* $\mathcal{C}$*.*

Let the trace of the toroid with the X_2*-*Z_2 *plane be the contour* $\mathcal{T}$ *of Fig. 3.3. It is most convenient to represent this contour with the aid of the non-dimensional variables u and v, which are defined as*

$$u \equiv \frac{x_2}{a}, \quad v \equiv \frac{z_2}{a} \tag{3.42}$$

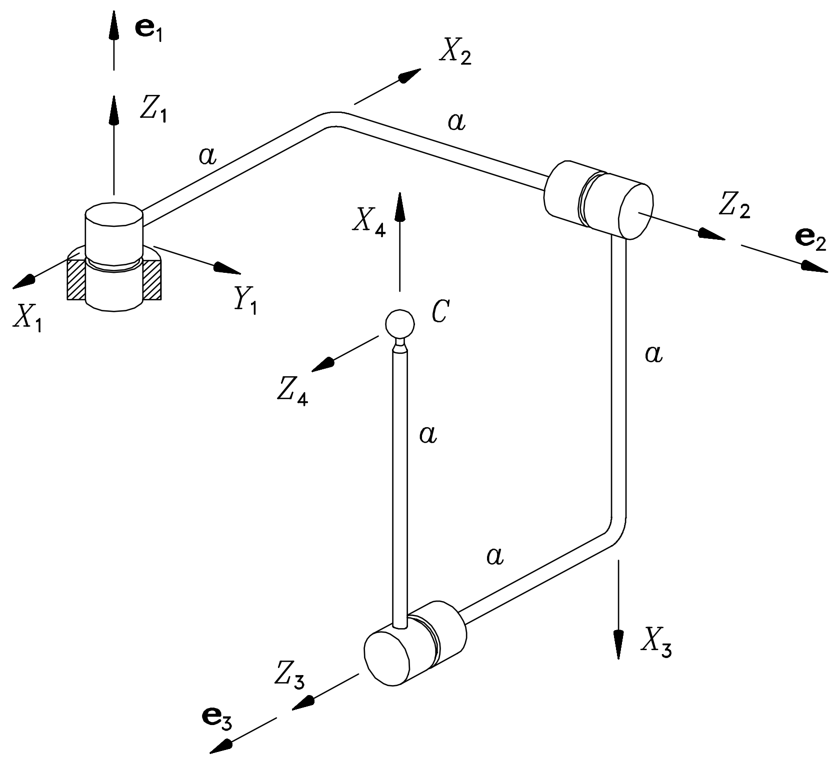

Figure 3.1: Manipulator configuration for $C(0,\, a,\, 0)$.

In terms of these new variables, the equation of $\mathcal{T}$ *becomes*

$$\mathcal{T}: \qquad h(u,v) \equiv u^4 + 2u^2v^2 + v^4 - 4u^2v - 4v^3 - 4u^2 + 4v^2 + 4 = 0 \tag{3.43}$$

The contour $\mathcal{T}$ *defined by the implicit function* $h(u,v) = 0$ *is displayed in Fig. 3.3.*

Now, the maximum distance r_M *of* O_1 *to* $\mathcal{T}$ *can be found as the solution of the optimization problem defined below:*

$$f(u,v) \equiv \frac{1}{2}[(u+1)^2 + v^2] \quad \rightarrow \quad \max_{u,v} \tag{3.44}$$

subject to eq.(3.43). We thus have an equality-constrained maximization problem. In order to find the normality conditions of this problem, we resort to Lagrange multipliers, thus defining a new, unconstrained, maximization problem:

$$F(u,v,\lambda) \equiv f + \lambda h \quad \rightarrow \quad \max_{u,v,\lambda} \tag{3.45}$$

The normality conditions of the foregoing problem are, thus,

$$\frac{\partial F}{\partial u} \equiv u + 1 + \lambda(4u^3 + 4uv^2 - 8uv - 8u) = 0 \tag{3.46a}$$

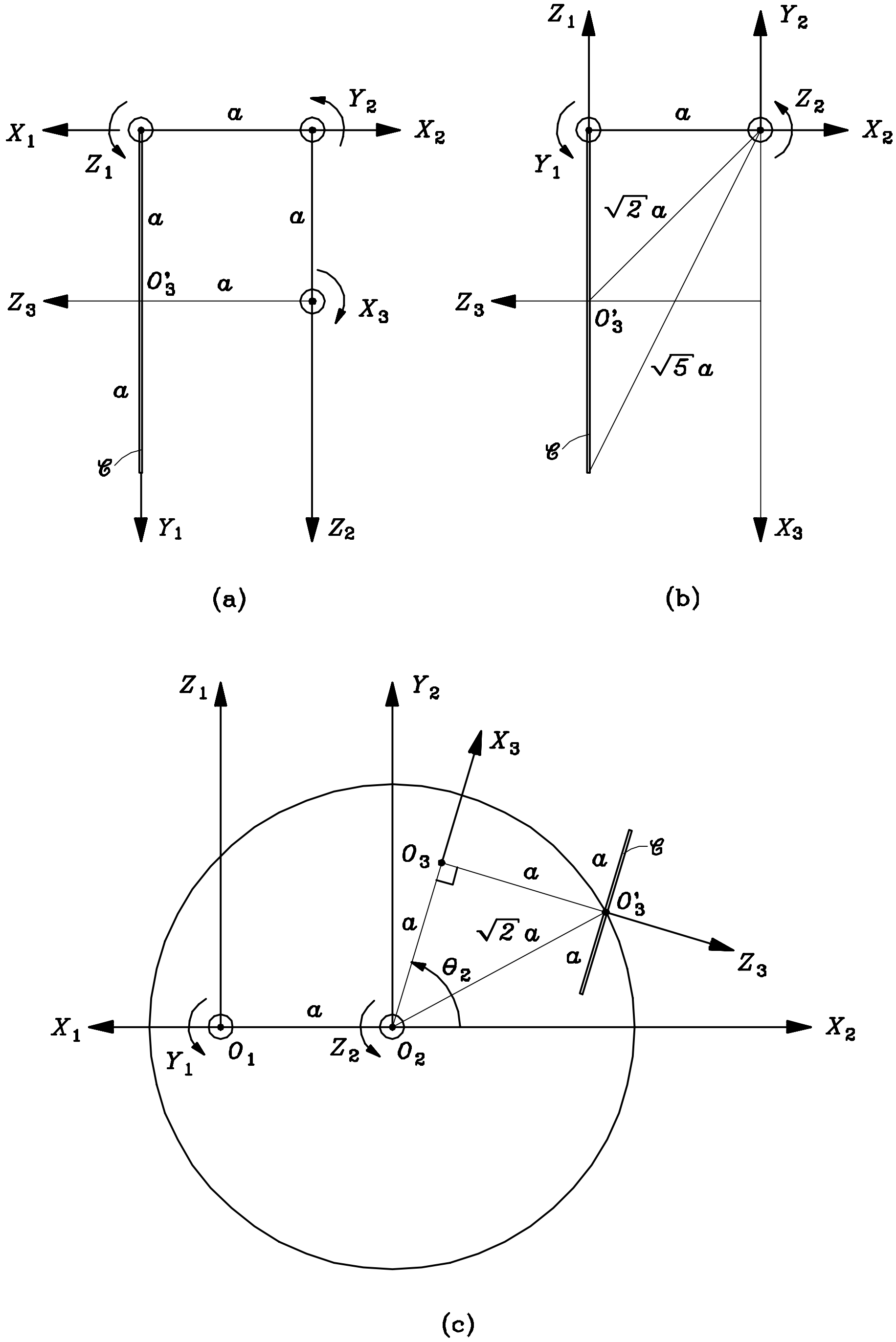

Figure 3.2: Side views of circle $\mathcal{C}$: (a) and (b) at the posture of Fig. 3.1; and (c) at an arbitrary posture for a given value of θ_2

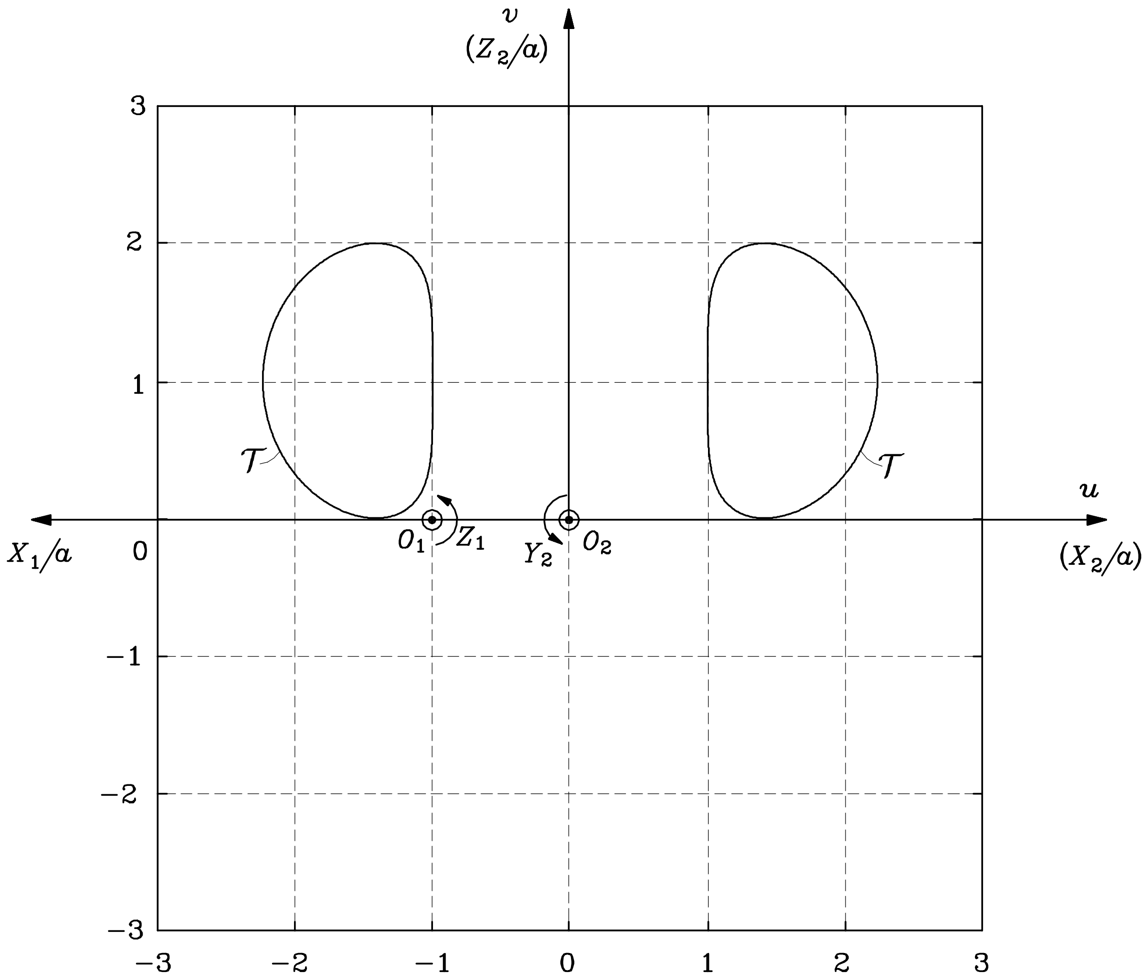

Figure 3.3: Contour of the trace $\mathcal{T}$ of the toroid on the u-v plane

$$\frac{\partial F}{\partial v} \equiv v + \lambda(4u^2v + 4v^3 - 4u^2 - 12v^2 + 8v) = 0 \qquad (3.46b)$$

$$\frac{\partial F}{\partial \lambda} \equiv u^4 + 2u^2v^2 + v^4 - 4u^2v - 4v^3 - 4u^2 + 4v^2 + 4 = 0 \qquad (3.46c)$$

the last equation being just a restatement of the constraint, eq.(3.43). Now we eliminate λ, the Lagrange multiplier, dialytically *(Salmon, 1964) from eqs.(3.46a & b). We do this by rewriting these two equations in linear homogeneous form in the "variables" λ and 1, namely,*

$$\begin{bmatrix} 4u^3 + 4uv^2 - 8uv - 8u & u+1 \\ 4u^2v + 4v^3 - 4u^2 - 12v^2 + 8v & v \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (3.47a)$$

Obviously, the foregoing equation requires a nontrivial solution—note that one component of the vector of "unknowns" is unity!—which in turn requires that the coefficient matrix be singular, i.e.,

$$\det\left(\begin{bmatrix} 4u^3 + 4uv^2 - 8uv - 8u & u+1 \\ 4u^2v + 4v^3 - 4u^2 - 12v^2 + 8v & v \end{bmatrix}\right) = 0 \qquad (3.47b)$$

Upon expansion,

$$4(u^3 + uv^2 - 2uv - 2u)v - 4(u^2v + v^3 - u^2 - 3v^2 + 2v)(u+1) = 0$$

or

$$\mathcal{S}: \quad u^3 - u^2v + uv^2 - v^3 + u^2 - 4uv + 3v^2 - 2v = 0 \qquad (3.47c)$$

Now, the maximum reach is found via the solution of the system of polynomial equations *(3.43) and (3.47c). The former is a quartic equation, the latter being cubic. The* Bezout number[2] *of the foregoing system of equations is defined as the product of the degrees of those equations, i.e., $4 \times 3 = 12$, which gives an upper bound of 12 for the number of solutions, both real and imaginary, of the problem at hand. One graphical means of obtaining estimates of the real solutions of this system consists in plotting the two corresponding contours in the u-v plane, as shown in Fig. 3.4. The maximum reach occurs apparently, at point A, of coordinates* (2.2, 1.4) *estimated by inspection, which leads to a visual estimate of r_M, namely,*

$$r_m \approx 3.5a \qquad (3.48)$$

[2]To define the Bezout number of a system of p polynomial equations in p variables x_1, x_2, ..., x_p, we look first at the ith equation: A typical term of this equation involves the product $x_1^{d_{1i}} x_2^{d_{2i}} \cdots x_p^{d_{pi}}$. The degree d_i of this equation is the maximum of $d_{1i} + d_{2i} + \ldots + d_{pi}$, for $i = 1, \ldots, N_i$, where N_i denotes the number of terms of the ith equation. The Bezout number N_B of this system is defined as $N_B = d_1 d_2 \ldots d_p$.

The four intersections of these two curves correspond to the four stationary values of the distance from a point in the trace $\mathcal{T}$ to the point O_1 in the u-v plane. Of these four intersections, two are local maxima and two local minima. The **normality** *of* ∇f*, which in this case is identical to the vector from O_1 to $\mathcal{T}$ at the intersection points, is to be highlighted.*

The foregoing system is solved more precisely using computer algebra, thus obtaining the four real solutions given below:

$$\begin{aligned} (u)_A &= 2.132242, & (v)_A &= 1.148990 \\ (u)_B &= -1.578095, & (v)_B &= 1.975316 \\ (u)_C &= -1.132242, & (v)_C &= 0.116796 \\ (u)_D &= 1.025308, & (v)_D &= 0.366325 \end{aligned}$$

which lead to reach values of

$$r_A = 3.459606a, \quad r_B = 2.058171a, \quad r_C = 0.176435a, \quad r_D = 2.058171a$$

for a global maximum reach of

$$r_M = 3.459606a$$

The value of a that will yield the foregoing maximum reach is thus found as

$$3.460a = 0.8772 \quad \Rightarrow \quad a = 0.2535 \text{ m}$$

thereby completing the solution.

The verification of the first- and second-order normality conditions is left as an exercise.

Example 3.3.2 (The Equilibrium Configuration of a Four-Link Chain)
We consider here the problem of determining the equilibrium configuration of a chain composed of four identical links of length L each, suspended at two points located at the same level, a distance d apart. This problem was proposed by Luenberger (1984) to illustrate methods of nonlinear programming. Here, we use a simplified version of this problem with the purpose of obtaining a solution by simple equation-solving.

At the outset, we exploit the symmetry of the problem, which enables us to reduce the number of design variables to only two, namely, the inclination of the two links on the left half of the chain. Let θ_i, for $i = 1, 2$, denote the angle made by the axis of the ith link from the vertical and μ denote the mass distribution per unit length,

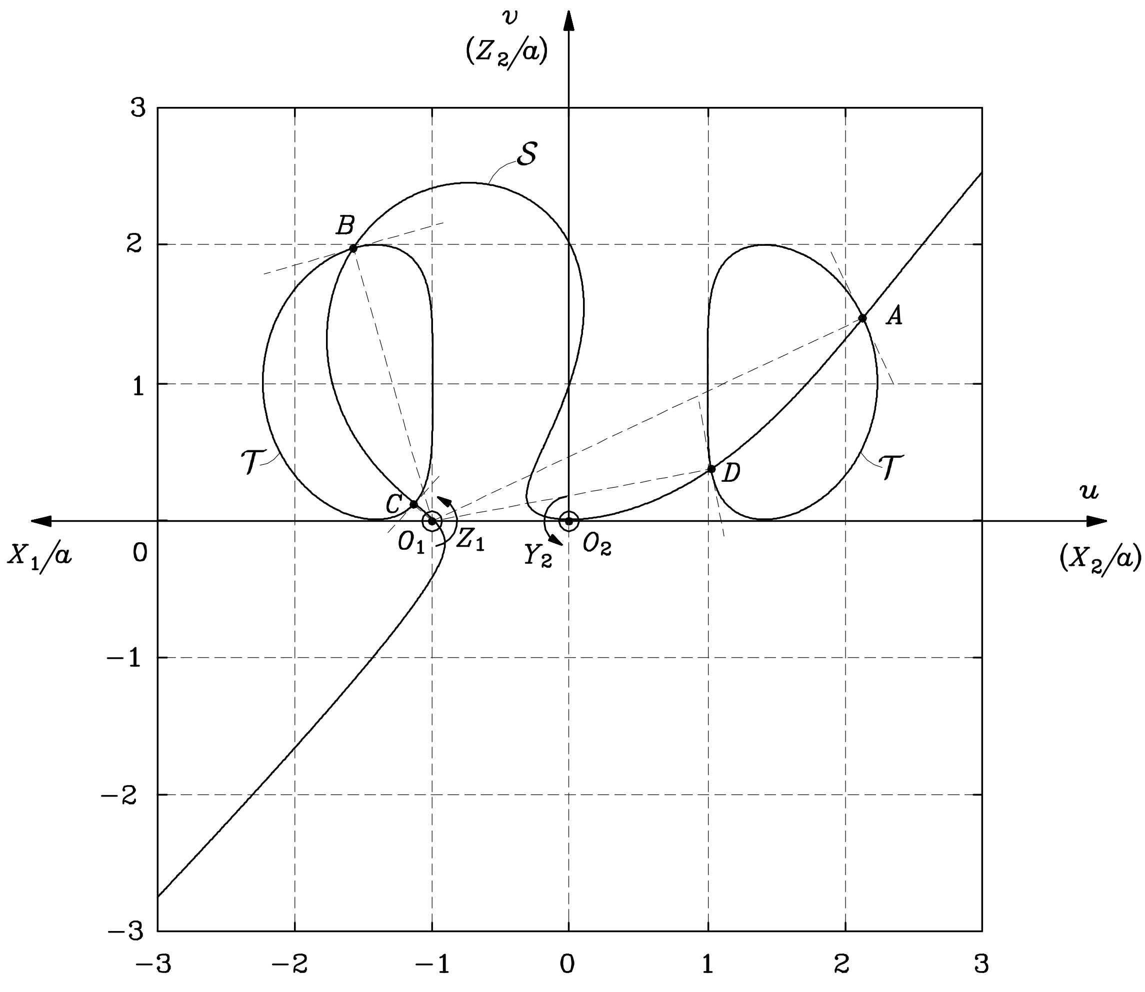

Figure 3.4: Plots of the two contours, $\mathcal{S}$ and $\mathcal{T}$

while g represent the gravity acceleration. The potential energy V of the whole chain is, thus, for an arbitrary configuration of the chain,

$$V(\theta_1, \theta_2) = -2\mu g L \left(\frac{1}{2}\cos\theta_1 + \cos\theta_1 + \frac{1}{2}\cos\theta_2\right)$$

which is a minimum at an equilibrium configuration. However, notice that the two design variables are not independent, for their horizontal span must be exactly $d/2$, i.e.,

$$L(\sin\theta_1 + \sin\theta_2) - \frac{d}{2} = 0$$

The optimum design problem at hand now has the form

$$f(\theta_1, \theta_2) \equiv \frac{V}{\mu g L} = -3\cos\theta_1 - \cos\theta_2 \quad \rightarrow \quad \min_{\theta_1, \theta_2}$$

subject to

$$h(\theta_1, \theta_2) = \sin\theta_1 + \sin\theta_2 - p = 0, \quad p \equiv \frac{d}{2L}$$

The Lagrangian of the problem is to be minimized, i.e.,

$$F(\theta_1, \theta_2) = -3\cos\theta_1 - \cos\theta_2 + \lambda(\sin\theta_1 + \sin\theta_2 - p) \quad \rightarrow \quad \min_{\theta_1, \theta_2, \lambda}$$

subject to no constraints. The normality conditions of the unconstrained problem are, thus

$$\frac{\partial F}{\partial \theta_1} = 3\sin\theta_1 + \lambda\cos\theta_1 = 0$$
$$\frac{\partial F}{\partial \theta_1} = \sin\theta_2 + \lambda\cos\theta_2 = 0$$
$$\frac{\partial F}{\partial \lambda} = \sin\theta_1 + \sin\theta_2 - p = 0$$

The problem has thus been reduced to solving the foregoing system of three nonlinear equations in three unknowns, θ_1, θ_2 and λ. While this nonlinear system can be solved using the Newton-Raphson method, the simplicity of the equations lends itself to a more comprehensive approach. Indeed, the Newton-Raphson method yields one single solution at a time, the user never knowing whether any other solutions exist. Moreover, there is no guarantee that the solution found is a minimum and not a maximum or a saddle point.

For starters, we can eliminate λ from the above equations, for it appears linearly in the first two of those. We thus rewrite those two equations in the form

$$\mathbf{A}\mathbf{x} = \mathbf{0}_2$$

with $\mathbf{0}_2$ *denoting the two-dimensional zero vector, while* $\mathbf{A}$ *and* $\mathbf{x}$ *are defined as*

$$\mathbf{A} \equiv \begin{bmatrix} \cos\theta_1 & 3\sin\theta_1 \\ \cos\theta_2 & \sin\theta_2 \end{bmatrix}, \quad \mathbf{x} \equiv \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \neq \mathbf{0}_2$$

Since the solution sought cannot be zero, the above homogeneous system must admit a nontrivial solution, which calls for $\mathbf{A}$ *to be singular, i.e.,*

$$\Delta(\theta_1, \theta_2) \equiv \det(\mathbf{A}) = 0$$

Upon expansion,

$$\Delta(\theta_1, \theta_2) = \cos\theta_1 \sin\theta_2 - 3\sin\theta_1 \cos\theta_2 = 0$$

which we shall call the reduced *normality condition. We thus have eliminated* λ dialytically *(Salmon, 1964), the problem thus reducing to a system of two equations in two unknowns,* $h(\theta_1, \theta_2) = 0$ *and* $\Delta(\theta_1, \theta_2) = 0$*. We can further reduce the same system to one single equation in one single unknown, which can be done by dialytic elimination as well. However, notice that dialytic elimination is applicable to systems of polynomial equations, while the two equations at hand are not polynomial; they are trigonometric. Nevertheless, by application of the well-known trigonometric "half-tan" identities:*

$$\cos x \equiv \frac{1^2_T}{1+T^2}, \quad \sin x \equiv \frac{2T}{1+T^2}, \quad T \equiv tan\left(\frac{x}{2}\right)$$

the two equations can be transformed into polynomial equations. We will not pursue here this elimination procedure. Instead, we plot the two foregoing functions in the θ_1-θ_2 *plane, the solutions sought being found visually at the intersection of the corresponding contours. In order to plot the contours, however, we must assign a numerical value to parameter* p*. By assuming* $d = 1.25$ m *and* $L = 0.5$ m*, we obtain* $p = 1.25$*. These contours are plotted in Fig. 3.5.*

The contours apparently intersect at two points, of coordinates estimated visually at

$$\theta_1 = 0.45, \quad \theta_2 = 1.00 \quad \text{and} \quad \theta_1 = 2.70, \quad \theta_2 = 2.20$$

These values are quite rough. Better values can be obtained by means of Newton-Raphson's method applied to the two nonlinear equations, using the foregoing estimates as initial guesses. Alternatively, the two equations can be solved dialytically by means of computer algebra. For example, upon invoking Maple's "`solve`*" procedure, the real roots below were reported:*

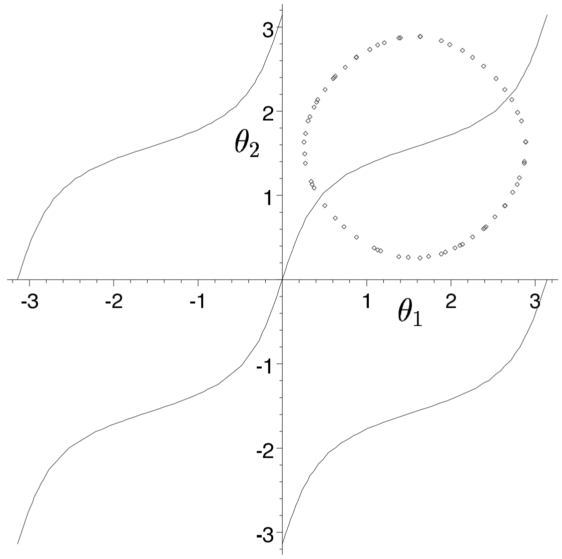

Figure 3.5: The reduced normality condition and the equality constraint (dashed contour)

$$\theta_1 = 0.4449420670,\ \theta_2 = 0.9607027573 \quad \text{and} \quad \theta_1 = 2.696650587,\ \theta_2 = 2.180889896$$

Translated into degrees, the foregoing angles read:

$$\theta_1 = 25.49330256°,\ \theta_2 = 55.04421335° \text{ and } \theta_1 = 154.5066974°,\ \theta_2 = 124.9557866°$$

The first solution corresponds, apparently, to a minimum, the second to a maximum. If this is the case, then the sum of the corresponding roots for the two solutions should be π*, which is the case. In fact, upon evaluation of the objective function at the two solutions, we obtain*

$$f(0.4449420670, 0.9607027573) = -1.640425478$$
$$f(2.696650587, 2.180889896) = 1.640425479$$

Notice the symmetry of the objective function at the two foregoing extrema.

The first- and second-order normality conditions should be verified numerically. The chain at its equilibrium configuration is displayed in Fig. 3.6.

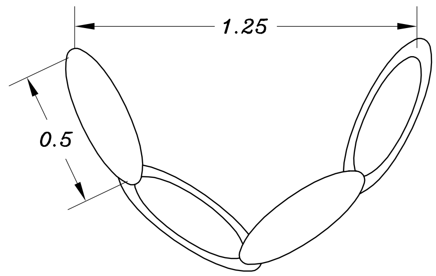

Figure 3.6: The equilibrium configuration of a four-link chain

3.4 Linear-Quadratic Problems

3.4.1 The Minimum-Norm Solution of Underdetermined Systems

We start by recalling a concept of paramount importance in optimization:

Definition 3.4.1 (Convex set) *A set of points $\mathcal{C}$ is* convex *if, given any two distinct points P_1 and P_2 of the set, then any point P of $\mathcal{C}$ comprised between P_1 and P_2 also belongs to the set. Otherwise, the set is* nonconvex.

More formally, if $\mathbf{x}_i$ denotes the position vector of P_i, for $i = 1, 2$, and $\mathbf{x}$ that of P, then, for any scalar α comprised in the interval $[\,0,\,1\,]$, we can express the position vector of P as a *convex combination* of those of P_1 and P_2, namely,

$$\mathbf{x} = \alpha\mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2 \tag{3.49}$$

We can thus rephrase the definition of convex set as

Definition 3.4.2 (Convex set—An alternative definition) *A set of points $\mathcal{C}$ is convex if, given any two distinct points of position vectors $\mathbf{x}_1$ and $\mathbf{x}_2$, then the point whose position vector is a convex combination of $\mathbf{x}_1$ and $\mathbf{x}_2$ also belongs to $\mathcal{C}$.*

Germane to the concept of convex set is that defined below:

Definition 3.4.3 (convex function) *A function $f(\mathbf{x})$ is convex if, for any $\mathbf{x}_1$ and $\mathbf{x}_2$, and a $\mathbf{x}$ defined as a convex combination of $\mathbf{x}_1$ and $\mathbf{x}_2$, and given, e.g., as in eq.(3.49),*

$$f(\mathbf{x}) \leq \alpha f(\mathbf{x}_1) + (1 - \alpha) f(\mathbf{x}_2) \tag{3.50}$$

Now we study the *underdetermined system* of linear equations

$$\mathbf{C}\mathbf{x} = \mathbf{d} \tag{3.51}$$

where $\mathbf{C}$ is a $p \times n$ matrix with $p < n$, all equations being assumed linearly independent. Apparently, the system *admits infinitely-many solutions.* Notice that the set of solutions of this equation does not form a vector space. Indeed, since $\mathbf{0}$ is not a solution, the solution set does not include the origin, which disqualifies the set from being a vector space. However, the same set has a quite interesting property:

Fact 3.4.1 *The set of solutions of the system (3.51) is convex.*

Proof: Assume that $\mathbf{x}_1$ and $\mathbf{x}_2$ are two distinct solutions of eq.(3.51), i.e.,

$$\mathbf{C}\mathbf{x}_1 = \mathbf{d} \tag{3.52a}$$

$$\mathbf{C}\mathbf{x}_2 = \mathbf{d} \tag{3.52b}$$

Now, for a real α such that $0 \leq \alpha \leq 1$, we have

$$\mathbf{C}(\alpha\mathbf{x}_1) = \alpha\mathbf{d} \tag{3.53a}$$

$$\mathbf{C}[(1-\alpha)\mathbf{x}_2] = (1-\alpha)\mathbf{d} \tag{3.53b}$$

Upon adding sidewise eqs.(3.53a & b), we obtain

$$\mathbf{C}[\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2] = \mathbf{d} \tag{3.54}$$

thereby completing the proof.

Geometrically, eq.(3.51) represents a plane embedded in n-dimensional space, offset from the origin. Each point of the plane thus has a position vector that is a solution. Out of the infinity of solutions satisfying the equation, then, there is one that lies closest to the origin. This is the *minimum-norm solution* of eq.(3.51). We derive below this solution upon solving the problem below:

$$f(\mathbf{x}) \equiv \frac{1}{2}\|\mathbf{x}\|^2 \quad \rightarrow \quad \min_{\mathbf{x}} \tag{3.55}$$

subject to eq.(3.51). As before, we transform the above constrained problem into an unconstrained one. We do this by means of Lagrange multipliers:

$$F(\mathbf{x}) \equiv f(\mathbf{x}) + \boldsymbol{\lambda}^T(\mathbf{C}\mathbf{x} - \mathbf{d}) \quad \rightarrow \quad \min_{\mathbf{x},\boldsymbol{\lambda}} \tag{3.56}$$

subject to no constraints. The normality conditions of this problem are, thus,

$$\frac{\partial F}{\partial \mathbf{x}} \equiv \nabla f + \mathbf{C}^T \boldsymbol{\lambda} = \mathbf{x} + \mathbf{C}^T \boldsymbol{\lambda} = \mathbf{0}_n \tag{3.57a}$$

$$\frac{\partial F}{\partial \boldsymbol{\lambda}} \equiv \mathbf{C}\mathbf{x} - \mathbf{d} = \mathbf{0}_p \tag{3.57b}$$

the second set of the above equations thus being just a restatement of the system of underdetermined equations (3.51). Solving for $\mathbf{x}$ from eq.(3.57a) yields

$$\mathbf{x} = -\mathbf{C}^T \boldsymbol{\lambda} \tag{3.58a}$$

which, when substituted into eq.(3.57b), leads to

$$-\mathbf{C}\mathbf{C}^T \mathbf{x} - \mathbf{d} = \mathbf{0}_p \tag{3.58b}$$

Since we assumed at the outset that the given eqs.(3.51) are linearly-independent, $\mathbf{C}$ is of full rank, and hence, the $p \times p$ symmetric matrix is nonsingular. As a result, this matrix is, in fact, positive-definite, the outcome being that eq.(3.58b) can be solved for $\mathbf{x}$ by means of the Cholesky decomposition. The result is, symbolically, the minimum-norm solution $\mathbf{x}_0$ sought:

$$\mathbf{x}_0 = \mathbf{C}^\dagger \mathbf{d} \tag{3.59a}$$

where

$$\mathbf{C}^\dagger = \mathbf{C}^T (\mathbf{C}\mathbf{C}^T)^{-1} \tag{3.59b}$$

which is the *right Moore-Penrose generalized inverse* of the rectangular matrix $\mathbf{C}$. One can see that the straightforward evaluation of $\mathbf{C}^\dagger$ by its definition, eq. (3.59a), involves the inversion of a matrix product, which is computationally costly and prone to ill conditioning, similar to the case of the right Moore-Penrose generalized inverse of eq.(2.38b). Moreover, the solution of eq. (3.59a) does not hold when $\mathbf{C}$ is rank-deficient.

An efficient and robust alternative to computing explicitly the right Moore-Penrose generalized inverse relies in Householder reflections, as explained below: First, a set of $n \times n$ Householder reflections $\mathbf{H}_1$, $\mathbf{H}_2$, $\ldots$, $\mathbf{H}_p$ is defined, such that the product $\mathbf{H} = \mathbf{H}_p \cdots \mathbf{H}_2 \mathbf{H}_1$ transforms $\mathbf{C}^T$ into upper-triangular form, thereby obtaining

$$\mathbf{H}\mathbf{C}^T = \begin{bmatrix} \mathbf{U} \\ \mathbf{O}_{n'p} \end{bmatrix} \tag{3.60}$$

where $\mathbf{U}$ is a $p \times p$ upper-triangular matrix, which is nonsingular because we have assumed that $\mathbf{C}$ is of full rank, while $\mathbf{O}_{n'p}$ is the $n' \times p$ zero matrix, with $n' \equiv n - p$.

Further, we rewrite eq.(3.51) in the form

$$\mathbf{C}\mathbf{H}^T\mathbf{H}\mathbf{x} = \mathbf{d} \tag{3.61}$$

which does not alter the original equation (3.51) because $\mathbf{H}$ is orthogonal. Letting $\mathbf{y} = \mathbf{H}\mathbf{x}$, from eqs. (3.60) and (3.61), one can realize that $\mathbf{x}$ and $\mathbf{y}$ have the same Euclidean norm, and hence, minimizing the norm of $\mathbf{y}$ is equivalent to minimizing that of $\mathbf{x}$. Thus, $\mathbf{x}$ will be the minimum-norm solution of the underdetermined system (3.51) if $\mathbf{y}$ is, correspondingly, the minimum-norm solution of the system

$$(\mathbf{H}\mathbf{C}^T)^T\mathbf{y} = \mathbf{d} \tag{3.62a}$$

Upon substitution of eq.(3.60) into eq.(3.62a), we obtain, with a suitable partitioning of $\mathbf{y}$,

$$[\,\mathbf{U}^T \quad \mathbf{O}_{n'p}^T\,]\begin{bmatrix}\mathbf{y}_U\\ \mathbf{y}_L\end{bmatrix} = \mathbf{d}, \quad \mathbf{y} \equiv \begin{bmatrix}\mathbf{y}_U\\ \mathbf{y}_L\end{bmatrix} \tag{3.62b}$$

which, upon expansion, leads to

$$\mathbf{U}^T\mathbf{y}_U + \mathbf{O}_{n'p}^T\mathbf{y}_L = \mathbf{d} \tag{3.62c}$$

whence it is apparent that $\mathbf{y}_L$ is undetermined, and hence, can be assigned *any* value, while $\mathbf{y}_U$ is determined because we have assumed that $\mathbf{C}$ is of full rank, $\mathbf{U}$ thus being nonsingular. If our intention is to minimize $\|\mathbf{x}\|$ or, equivalently, $\|\mathbf{y}\|$, whose square is given by

$$\|\mathbf{y}\|^2 = \|\mathbf{y}_U\|^2 + \|\mathbf{y}_L\|^2$$

it is apparent that the optimum choice of $\mathbf{y}_L$ is $\mathbf{y}_L = \mathbf{0}_{n'}$, with $\mathbf{0}_{n'}$ denoting the $(n-p)$-dimensional zero vector. Therefore, the minimum-norm solution $\mathbf{y}_0$ of eq.(3.62a) takes on the form:

$$\mathbf{y}_0 = \begin{bmatrix}\mathbf{U}^{-T}\mathbf{d}\\ \mathbf{0}_{n'}\end{bmatrix} \tag{3.63}$$

i.e., the last $(n-p)$ components of $\mathbf{y}_0$ are zero. In this way, $\mathbf{y}_0$ verifies eq. (3.61) and has a minimum norm. Then, the minimum-norm solution $\mathbf{x}_0$ can be readily computed as

$$\mathbf{x}_0 = \mathbf{H}^T\mathbf{y}_0 \tag{3.64}$$

The Case of a Rank-Deficient C Matrix

If $\mathbf{C}$ is rank-deficient, with $\text{rank}(\mathbf{C}) = r < p$, then we can proceed as described above with only r Householder reflections, namely, $\mathbf{H} = \mathbf{H}_r\mathbf{H}_{r-1}\dots\mathbf{H}_1$, such that

$$\mathbf{H}\mathbf{C}^T = \begin{bmatrix}\overline{\mathbf{U}}\\ \overline{\mathbf{O}}\end{bmatrix}, \quad \mathbf{y} \equiv \mathbf{H}\mathbf{x} \tag{3.65}$$

where $\overline{\mathbf{U}}$ is a full-rank $r \times p$ matrix with zero entries in its lower-left "corner"—this matrix has an upper-trapezoidal form—and $\overline{\mathbf{O}}$ defined as the $(n-r) \times p$ zero matrix.

Note that, in general, the rank of $\mathbf{C}$ is not known in advance. It is first learned when the p Householder reflections of Subsection 3.4.1 are defined to bring $\mathbf{C}^T$ into upper-triangular form. In the presence of a rank-deficient matrix $\mathbf{C}$ of rank $r < p$, The last $n - r$ rows of $\mathbf{HC}^T$ are all zero, and the last $p - r$ $\mathbf{H}_i$ matrices are all identical.

Upon application of the foregoing r Householder reflections, eq. (3.62b) becomes

$$[\overline{\mathbf{U}}^T \quad \overline{\mathbf{O}}] \begin{bmatrix} \overline{\mathbf{y}}_U \\ \overline{\mathbf{y}}_L \end{bmatrix} = \mathbf{d}, \quad \mathbf{y} \equiv \begin{bmatrix} \overline{\mathbf{y}}_U \\ \overline{\mathbf{y}}_L \end{bmatrix} \tag{3.66}$$

whence,

$$\overline{\mathbf{U}}^T \overline{\mathbf{y}}_U + \overline{\mathbf{O}} \overline{\mathbf{y}}_L = \mathbf{d} \tag{3.67}$$

Apparently, $\|\mathbf{x}\| = \|\mathbf{y}\|$, and hence, minimizing one norm, one minimizes the other one. Moreover,

$$\|\mathbf{y}\|^2 = \|\overline{\mathbf{y}}_U\|^2 + \|\overline{\mathbf{y}}_L\|^2$$

Therefore, the optimum choice of $\mathbf{y}$ is the one for which $\overline{\mathbf{y}}_L = \mathbf{0}_{n''}$, with $\mathbf{0}_{n''}$ denoting the n''-dimensional zero vector, and $n'' \equiv n - r$, eq.(3.67) thus reducing to

$$\overline{\mathbf{U}}^T \overline{\mathbf{y}}_U = \mathbf{d} \tag{3.68}$$

where $\overline{\mathbf{U}}^T$ is a $p \times r$ matrix with zero entries in its upper corner, i.e., this matrix has the form

$$\overline{\mathbf{U}}^T = \begin{bmatrix} \overline{\mathbf{L}} \\ \mathbf{M} \end{bmatrix} \tag{3.69}$$

in which $\overline{\mathbf{L}}$ is a nonsingular $r \times r$ lower-triangular matrix and $\mathbf{M}$ is a $(p - r) \times r$ matrix. Moreover, since $\overline{\mathbf{U}}$ has been assumed of full rank, $\overline{\mathbf{U}}^T$ is also of full rank, its last $p - r$ rows being linearly dependent from its first r rows. That is, the $p - r$ rows of $\mathbf{M}$ are linearly dependent from the r rows of $\overline{\mathbf{L}}$. This means that $\overline{\mathbf{y}}_U$ is determined from the first r equations of eq.(3.68). We can thus use only those equations, which are, moreover, in lower-triangular form already, to compute $\mathbf{y}_U$ by *forward substitution*. *Symbolically*, then, we have

$$\overline{\mathbf{y}}_U = \overline{\mathbf{L}}^{-1} \mathbf{d}, \quad \mathbf{x}_0 = \mathbf{H}^T \begin{bmatrix} \overline{\mathbf{L}}^{-1} \mathbf{d} \\ \mathbf{0}_{n''} \end{bmatrix} \tag{3.70}$$

Alternatively, and if CPU time is not an issue, we can use all redundant scalar equations of that vector equation. We do this, then, by application of another set of r Householder reflections, $\overline{\mathbf{H}}_1$, $\overline{\mathbf{H}}_2$, ..., $\overline{\mathbf{H}}_r$, thereby obtaining

$$\overline{\mathbf{H}}\mathbf{U}^T\mathbf{y}_U = \overline{\mathbf{H}}\mathbf{d}, \quad \overline{\mathbf{H}} \equiv \overline{\mathbf{H}}_1\overline{\mathbf{H}}_2 \ldots \overline{\mathbf{H}}_r \tag{3.71}$$

whence the optimum solution is obtained in the way explained for overdetermined systems in Section 2.4. The details are left as an exercise.

Example 3.4.1 (The Solution of $\mathbf{a} \times \mathbf{x} = \mathbf{b}$)
Let $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{x}$ *be three 3-dimensional Cartesian vectors. We would like to solve the equation*

$$\mathbf{a} \times \mathbf{x} = \mathbf{b}$$

for $\mathbf{x}$. *It is well known, however, that the foregoing equation contains only two independent scalar equations, which prevents us from finding "the* $\mathbf{x}$*" that verifies that equation. Thus, we can proceed by finding a specific* $\mathbf{x}$, $\mathbf{x}_0$, *that verifies* any *two of these three equations and that is of minimum norm. To this end, we expand that equation into its three components:*

$$a_2x_3 - a_3x_2 = b_1$$
$$a_3x_1 - a_1x_3 = b_2$$
$$a_1x_2 - a_2x_1 = b_3$$

Note that the foregoing equation can be cast in the form of eq.(2.3) if we define matrix $\mathbf{A}$ *as*

$$\mathbf{A} \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

which is apparently skew-symmetric, i.e.,

$$\mathbf{A}^T = -\mathbf{A}$$

In fact, $\mathbf{A}$ *is the cross-product matrix of* $\mathbf{a}$. *Picking up, for example, the first two scalar equations above, we obtain an underdetermined system of the form (3.51), with*

$$\mathbf{C} \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

and hence, the corresponding minimum-norm solution $\mathbf{x}_0$ *is given by eqs.(3.59a & b), with*

$$\mathbf{C}\mathbf{C}^T = \begin{bmatrix} a_2^2 + a_3^2 & -a_1a_2 \\ -a_1a_2 & a_1^2 + a_3^2 \end{bmatrix}$$

Hence,

$$(\mathbf{C}\mathbf{C}^T)^{-1} = \frac{1}{\Delta}\begin{bmatrix} a_1^2 + a_3^2 & a_1 a_2 \\ a_1 a_2 & a_2^2 + a_3^2 \end{bmatrix}$$

where

$$\Delta \equiv \det(\mathbf{C}\mathbf{C}^T) = (a_2^2 + a_3^2)(a_1^2 + a_3^2) - a_1^2 a_2^2 > 0$$

a relation that the reader can readily prove. Therefore,

$$\mathbf{C}^\dagger = \frac{1}{\Delta}\begin{bmatrix} a_1 a_2 a_3 & (a_2^2 + a_3^2) a_3 \\ -(a_1^2 + a_3^2) a_3 & -a_1 a_2 a_3 \\ a_2 a_3^2 & -a_1 a_3^2 \end{bmatrix}$$

and

$$\mathbf{x}_0 = \frac{1}{\Delta}\begin{bmatrix} a_1 a_2 a_3 b_1 + (a_2^2 + a_3^2) a_3 b_2 \\ -(a_1^2 + a_3^2) a_3 b_1 - a_1 a_2 a_3 b_2 \\ a_2 a_3^2 b_1 - a_1 a_3^2 b_2 \end{bmatrix}$$

Notice that the foregoing solution depends on the condition $a_3 \neq 0$*. If* $a_3 = 0$*, or very close to 0, then* $\mathbf{C}$ *becomes either ill-conditioned or rank-deficient, which is bad news. Apparently, the foregoing solution has an element of arbitrariness that may lead either to ill-conditioning or to rank-deficiency. There is no guarantee that the two equations chosen are the best choice from the condition-number viewpoint. Besides, that approach leaves aside useful information, that of the deleted equation. The alternative approach uses all three equations, to which one fourth equation is adjoined, namely, the minimum-norm condition, as described below.*

First we observe that, if $\mathbf{x}$ *has been found that verifies the given cross-product equation, then any other vector* $\mathbf{x}+\alpha\mathbf{a}$*, for* $\alpha \in \mathbb{R}$*, verifies that equation. Apparently, then, the minimum-norm* $\mathbf{x}$ *is that whose component along* $\mathbf{a}$ *vanishes, i.e.,*

$$\mathbf{a}^T\mathbf{x} = 0$$

Upon adjoining the foregoing equation to the original three, we end up with an apparently overdetermined system of four equations with three unknowns, of the form

$$\mathbf{M}\mathbf{x} = \mathbf{n}$$

where $\mathbf{M}$ *and* $\mathbf{n}$ *are given by*

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} \\ \mathbf{a}^T \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix}$$

Hence, $\mathbf{M}$ *is a* 4×3 *matrix, while* $\mathbf{n}$ *is a 4-dimensional vector. The least-square approximation of the new system is, then, the minimum-norm solution of the original* *system, provided the latter is verified exactly, which it is, as will become apparent. Indeed, the least-square approximation of the new system takes the form*

$$\mathbf{x}_L = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{n} \tag{3.73}$$

While we have strongly advised against the explicit computation of generalized inverses, our advice is valid only as pertaining to numerical computations. In the case at hand, we will pursue not a numerical, but rather a symbolic *computation of the solution sought.*

The first issue now is whether $\mathbf{M}^T\mathbf{M}$ *is invertible, but it is so and, moreover, its inverse is extremely simple to find:*

$$\mathbf{M}^T\mathbf{M} = [\,\mathbf{A}^T \quad \mathbf{a}\,]\begin{bmatrix}\mathbf{A}\\ \mathbf{a}^T\end{bmatrix} = \mathbf{A}^T\mathbf{A} + \mathbf{a}\mathbf{a}^T$$

But, since $\mathbf{A}$ *is skew-symmetric,*

$$\mathbf{M}^T\mathbf{M} = -\mathbf{A}^2 + \mathbf{a}\mathbf{a}^T$$

As the reader can readily verify, moreover,

$$\mathbf{A}^2 = -\|\mathbf{a}\|^2\mathbf{1} + \mathbf{a}\mathbf{a}^T$$

Hence,

$$\mathbf{M}^T\mathbf{M} = \|\mathbf{a}\|^2\mathbf{1}$$

which means that $\mathbf{M}$ *is* isotropic, *i.e., optimally-conditioned. Therefore,*

$$(\mathbf{M}^T\mathbf{M})^{-1} = \frac{1}{\|\mathbf{a}\|^2}\mathbf{1}$$

That is ,

$$\mathbf{x}_L = \frac{1}{\|\mathbf{a}\|^2}\mathbf{1}\,[\,\mathbf{A}^T \quad \mathbf{a}\,]\begin{bmatrix}\mathbf{b}\\ 0\end{bmatrix} = \frac{1}{\|\mathbf{a}\|^2}\mathbf{A}^T\mathbf{b}$$

which can be further expressed as

$$\mathbf{x}_L = -\frac{\mathbf{a}\times\mathbf{b}}{\|\mathbf{a}\|^2} \tag{3.74}$$

thereby obtaining a much simpler expression than that displayed above as $\mathbf{x}_0$.

3.4.2 Least-Square Problems Subject to Linear Constraints

Given the system of linear equations

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3.75}$$

where $\mathbf{A}$ is a full-rank $q \times n$ matrix, with $q > n$, and $\mathbf{b}$ is a q-dimensional vector, find an n-dimensional vector $\mathbf{x}$ that verifies the above system with the least-square error, subject to the linear equality constraints

$$\mathbf{C}\mathbf{x} = \mathbf{d} \tag{3.76}$$

with $\mathbf{C}$ a full-rank $p \times n$ matrix and $\mathbf{d}$ a p-dimensional vector. Moreover, $\mathbf{W}$ is a $q \times q$ positive-definite weighting matrix, with q, p and n subject to

$$q + p > n \quad \text{and} \quad n > p \tag{3.77}$$

The least-square error of eqs.(3.75) is defined as

$$f \equiv \frac{1}{2}(\mathbf{A}\mathbf{x} - \mathbf{b})^T \mathbf{W}(\mathbf{A}\mathbf{x} - \mathbf{b}) \tag{3.78}$$

As usual, we solve this problem by introducing Lagrange multipliers:

$$F(\mathbf{x}; \boldsymbol{\lambda}) \equiv f(\mathbf{x}) + \boldsymbol{\lambda}^T(\mathbf{C}\mathbf{x} - \mathbf{d}) \quad \min_{\mathbf{x}, \boldsymbol{\lambda}} \tag{3.79}$$

subject to no constraints.

The first-order normality conditions of the foregoing problem are

$$\frac{\partial F}{\partial \mathbf{x}} \equiv \mathbf{A}^T\mathbf{W}(\mathbf{A}\mathbf{x} - \mathbf{b}) + \mathbf{C}^T\boldsymbol{\lambda} = \mathbf{0}_n \tag{3.80a}$$

$$\frac{\partial F}{\partial \boldsymbol{\lambda}} \equiv \mathbf{C}\mathbf{x} - \mathbf{d} = \mathbf{0}_p \tag{3.80b}$$

Since $\mathbf{A}$ is assumed of full rank and $\mathbf{W}$ is positive-definite, we can solve eq.(3.80a) for $\mathbf{x}$ in terms of $\boldsymbol{\lambda}$, namely,

$$\mathbf{x} = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{W}\mathbf{b} - \mathbf{C}^T\boldsymbol{\lambda}) \tag{3.81}$$

Upon substituting the above expression into eq.(3.80b), we obtain

$$\mathbf{C}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{C}^T\boldsymbol{\lambda} = \mathbf{C}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\mathbf{b} - \mathbf{d}$$

whence,

$$\boldsymbol{\lambda} = [\mathbf{C}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{C}^T]^{-1}[\mathbf{C}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\mathbf{b} - \mathbf{d}] \tag{3.82}$$

Now, the foregoing expression for $\boldsymbol{\lambda}$ is substituted, in turn, into eq.(3.81), thereby obtaining the optimum value of $\mathbf{x}$, $\mathbf{x}_0$, namely,

$$\mathbf{x}_0 = \mathbf{PQb} + \mathbf{Rd} \tag{3.83a}$$

where $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are the $n \times n$-, $n \times m$- and $n \times p$ matrices given below:

$$\mathbf{P} = \mathbf{1}_n - \mathbf{RC} \tag{3.83b}$$

$$\mathbf{Q} = (\mathbf{A}^T\mathbf{WA})^{-1}\mathbf{A}^T\mathbf{W} \tag{3.83c}$$

$$\mathbf{R} = (\mathbf{A}^T\mathbf{WA})^{-1}\mathbf{C}^T[\mathbf{C}(\mathbf{A}^T\mathbf{WA})^{-1}\mathbf{C}^T]^{-1} \tag{3.83d}$$

with $\mathbf{1}_n$ standing for the $n \times n$ identity matrix. The solution derived above, while being exact, for it is symbolic, is unsuitable for numerical implementation. Indeed, this solution contains inversions of products of several matrices times their transposes, which brings about ill-conditioning. Various approaches to the numerical solution of this problem will be studied in Ch. 4.

3.5 Equality-Constrained Nonlinear Least Squares

We consider here the problem of finding the least-square error f of an overdetermined system of nonlinear equations, namely,

$$\boldsymbol{\phi}(\mathbf{x}) = \mathbf{0} \tag{3.84a}$$

subject to the nonlinear constraints

$$\mathbf{h}(\mathbf{x}) = \mathbf{0} \tag{3.84b}$$

In general, moreover, the various scalar equations of eq.(3.84a) have different relevance and are, hence, assigned different weights, which then leads to a problem of *weighted least squares*, namely,

$$f(\mathbf{x}) = \frac{1}{2}\boldsymbol{\phi}^T\mathbf{W}\boldsymbol{\phi} \quad \rightarrow \quad \min_{\mathbf{x}} \tag{3.85}$$

subject to eq.(3.84b).

In the foregoing problem, $\boldsymbol{\phi}$ and $\mathbf{x}$ are q- and n-dimensional vectors, respectively, with $q > n$, and $\mathbf{W}$ is a $q \times q$ positive-definite weighting matrix. Moreover, $\mathbf{h}$ is a l-dimensional vector of nonlinear constraints.

The normality conditions of the problem at hand are derived directly from those of the general equality-constrained problem, namely, eq.(3.27) or its dual counterpart, eq.(3.35). In our case,

$$\nabla f = \left(\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}\right)^T \frac{\partial f}{\partial \boldsymbol{\phi}} \tag{3.86a}$$

where

$$\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}} \equiv \boldsymbol{\Phi}(\mathbf{x}), \quad \frac{\partial f}{\partial \boldsymbol{\phi}} = \mathbf{W}\boldsymbol{\phi}(\mathbf{x}) \tag{3.86b}$$

i.e., $\boldsymbol{\Phi}(\mathbf{x})$ denotes the *Jacobian matrix* of $\boldsymbol{\phi}(\mathbf{x})$ with respect to $\mathbf{x}$. Hence,

$$\nabla f = \boldsymbol{\Phi}^T \mathbf{W} \boldsymbol{\phi} \tag{3.86c}$$

where we have dispensed with the argument $\mathbf{x}$ for simplicity.

The normality condition (3.27) thus reduces to

$$[\mathbf{1} - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}]\boldsymbol{\Phi}^T\mathbf{W}\boldsymbol{\phi} = \mathbf{0}_n \tag{3.87}$$

What this condition states is that, at a stationary point, $\boldsymbol{\phi}$, or $\boldsymbol{\Phi}^T\mathbf{W}\boldsymbol{\phi}$ for that matter, need not vanish; only the projection of the latter onto the nullspace of the gradient of the constraints must vanish.

The dual form of the same normality conditions, in turn, reduces to

$$\mathbf{L}^T\boldsymbol{\Phi}^T\mathbf{W}\boldsymbol{\phi} = \mathbf{0}_{n'} \tag{3.88}$$

with $\mathbf{L}$ indicating an $n \times (n-l)$ orthogonal complement of $\mathbf{J}$, as defined in eq.(3.33).

The second-order normality conditions are now derived by assuming that we have found a stationary value of the design-variable vector, $\mathbf{x}_0$. This means that

$$\mathbf{L}^T\boldsymbol{\Phi}^T(\mathbf{x}_0)\mathbf{W}\boldsymbol{\phi}(\mathbf{x}_0) = \mathbf{0}_{n'} \tag{3.89}$$

Next, we impose the condition that, for any *feasible move* $\Delta\mathbf{x}$, the corresponding increment of f, Δf, be positive. A feasible move is obtained by resorting to the orthogonal complement $\mathbf{L}$, namely,

$$\Delta\mathbf{x} = \mathbf{L}\Delta\mathbf{u} \tag{3.90}$$

On the other hand, upon expansion of $\boldsymbol{\phi}(\mathbf{x}_0 + \Delta\mathbf{x})$ to a first order,

$$\begin{aligned} f(\mathbf{x}_0 + \Delta\mathbf{x}) &\approx \frac{1}{2}(\boldsymbol{\phi} + \Delta\boldsymbol{\phi})^T\mathbf{W}(\boldsymbol{\phi} + \Delta\boldsymbol{\phi}) \\ &= \underbrace{\frac{1}{2}\boldsymbol{\phi}^T\mathbf{W}\boldsymbol{\phi}}_{f(\mathbf{x}_0)} + \frac{1}{2}(\boldsymbol{\phi}^T\mathbf{W}\Delta\boldsymbol{\phi} + \Delta\boldsymbol{\phi}^T\mathbf{W}\boldsymbol{\phi}) + \frac{1}{2}\Delta\boldsymbol{\phi}^T\mathbf{W}\Delta\boldsymbol{\phi} \end{aligned}$$

whence,

$$\Delta f \equiv f(\mathbf{x}_0 + \Delta\mathbf{x}) - f(\mathbf{x}_0) \approx \frac{1}{2}(\boldsymbol{\phi}^T\mathbf{W}\Delta\boldsymbol{\phi} + \Delta\boldsymbol{\phi}^T\mathbf{W}\boldsymbol{\phi}) + \frac{1}{2}\Delta\boldsymbol{\phi}^T\mathbf{W}\Delta\boldsymbol{\phi} \tag{3.91a}$$

where

$$\Delta\boldsymbol{\phi} = \boldsymbol{\Phi}\Delta\mathbf{x} = \boldsymbol{\Phi}\mathbf{L}\Delta\mathbf{u} \tag{3.91b}$$

Therefore,

$$\begin{aligned}\Delta f &= \frac{1}{2}[\underbrace{\boldsymbol{\phi}^T(\mathbf{x}_0)\mathbf{W}\boldsymbol{\Phi}(\mathbf{x}_0)\mathbf{L}}_{=\mathbf{0}^T_{n-l}}\Delta\mathbf{u} + \Delta\mathbf{u}^T\underbrace{\mathbf{L}^T\boldsymbol{\Phi}^T(\mathbf{x}_0)\mathbf{W}\boldsymbol{\phi}(\mathbf{x}_0)}_{=\mathbf{0}_{n-l}}] \\ &\quad + \frac{1}{2}\Delta\mathbf{u}^T\mathbf{L}^T\boldsymbol{\Phi}^T(\mathbf{x}_0)\mathbf{W}\boldsymbol{\Phi}\mathbf{L}\Delta\mathbf{u} \\ &= \frac{1}{2}\Delta\mathbf{u}^T\mathbf{L}^T\boldsymbol{\Phi}^T(\mathbf{x}_0)\mathbf{W}\boldsymbol{\Phi}(\mathbf{x}_0)\mathbf{L}\Delta\mathbf{u} > 0 \end{aligned} \tag{3.91c}$$

the terms inside the brackets vanishing by virtue of the first-order normality conditions. What conditions (3.91c) state is that, for the stationary value $\mathbf{x}_0$ to be a minimum, the feasible Hessian $\mathbf{L}^T\boldsymbol{\Phi}^T(\mathbf{x}_0)\mathbf{W}\boldsymbol{\Phi}(\mathbf{x}_0)\mathbf{L}$ must be positive-definite. However, at the outset we defined $\mathbf{W}$ as positive-definite, while $\mathbf{L}$ is of full rank. Hence, the feasible Hessian is necessarily positive-definite, except for points at which $\boldsymbol{\Phi}$ becomes rank-deficient, at which the said Hessian becomes positive-semidefinite. As a consequence, then,

Fact 3.5.1 *To a first-order approximation of $\boldsymbol{\phi}(\mathbf{x})$, a stationary point of the weighted least-square approximation of the overdetermined nonlinear system $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{0}$ is a minimum; never a maximum or a saddle point.*

3.6 Linear Least-Square Problems Under Quadratic Constraints

An important family of design problems lends itself to a formulation whereby the objective function is quadratic in a linear function of the design vector $\mathbf{x}$, while the constraints are quadratic in $\mathbf{x}$. Contrary to the case of linear least-squares subject to linear constraints, this family of problems does not allow, in general, for closed-form solutions, the reason being that their normal equations are nonlinear. Let us consider

$$f(\mathbf{x}) \equiv \frac{1}{2}(\mathbf{b} - \mathbf{A}\mathbf{x})^T\mathbf{W}(\mathbf{b} - \mathbf{A}\mathbf{x}) \quad \rightarrow \quad \min_{\mathbf{x}} \tag{3.92a}$$

subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}_l \tag{3.92b}$$

where $\mathbf{A}$ is a $q \times n$ full-rank matrix, with $q > n$, $\mathbf{W}$ is a $q \times q$ positive-definite weighting matrix, while $\mathbf{h}$, $\mathbf{x}$ and $\mathbf{b}$ are l-, n- and q-dimensional vectors, respectively, with

$$q + l > n, \quad n > l \tag{3.92c}$$

Moreover, in this particular case, the ith component of vector $\mathbf{h}$ is quadratic, namely,

$$h_i(\mathbf{x}) \equiv \frac{1}{2}\mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i + r_i \tag{3.92d}$$

in which $\mathbf{P}_i$ is a known $n \times n$ symmetric matrix, while $\mathbf{q}_i$ is a n-dimensional given vector and r_i is a given scalar. Apparently, then, the ith row of $\mathbf{J}$, the Jacobian of $\mathbf{h}$ with respect to $\mathbf{x}$, takes the form

$$\frac{\partial h_i}{\partial \mathbf{x}} = \mathbf{P}_i\mathbf{x} + \mathbf{q} \tag{3.93}$$

whence $\mathbf{J}$ is *linear* in $\mathbf{x}$. The first-order normality conditions (3.27) now take the form

$$[\mathbf{1}_n - \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}\mathbf{J}]\mathbf{A}^T\mathbf{W}(\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{0}_n \tag{3.94}$$

It is thus apparent that, although $\mathbf{J}$ is linear in $\mathbf{x}$, the normality conditions are polynomial, thereby leading to a problem lacking a closed-form solution, except for special cases, like the one included below.

Example 3.6.1 (A Quadratic Objective Function with a Quadratic Constraint)

Consider an optimization problem with an objective function defined as

$$f(\mathbf{x}) = \frac{1}{2}(9x_1^2 - 8x_1x_2 + 3x_2^2) \quad \rightarrow \quad \min_{x_1, x_2}$$

subject to the quadratic condition

$$h(\mathbf{x}) = x_1^2 + x_2^2 - 1 = 0 \tag{3.95}$$

The objective function $f(\mathbf{x})$ *can be factored as*

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{b} - \mathbf{A}\mathbf{x})^T\mathbf{W}(\mathbf{b} - \mathbf{A}\mathbf{x})$$

with

$$\mathbf{A} = \mathbf{1}_2, \quad \mathbf{b} = \mathbf{0}_2, \quad \mathbf{W} = \begin{bmatrix} 9 & -4 \\ -4 & 3 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{3.96}$$

Solution: *Upon adjoining the constraint to the objective function, we obtain the Lagrangian* $F(\mathbf{x}; \boldsymbol{\lambda})$*, namely,*

$$F(\mathbf{x}; \boldsymbol{\lambda}) = \frac{1}{2}(9x_1^2 - 8x_1x_2 + 3x_2^2) + \lambda(x_1^2 + x_2^2 - 1)$$

that we want to minimize subject to no constraints. The normality conditions are, thus

$$\begin{aligned} \frac{\partial F}{\partial x_1} &= 9x_1 - 4x_2 + 2\lambda x_1 = 0 \\ \frac{\partial F}{\partial x_2} &= -4x_1 + 3x_2 + 2\lambda x_2 = 0 \\ \frac{\partial F}{\partial \lambda} &= x_1^2 + x_2^2 - 1 = 0 \end{aligned}$$

We can now eliminate λ *from the first and the second of the above equations. We do this dialytically, i.e., we write these two equations in linear homogeneous form in* λ *and 1, i.e.,*

$$\mathbf{M}\mathbf{y} = \mathbf{0}_2$$

where

$$\mathbf{M} = \begin{bmatrix} 2x_1 & 9x_1 - 4x_2 \\ 2x_2 & -4x_1 + 3x_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \neq \mathbf{0}_2$$

with $\mathbf{0}_2$ *denoting the 2-dimensional zero vector. Now, the above linear homogeneous equation in* $\mathbf{y}$ *cannot be zero, for* $\mathbf{y} \neq \mathbf{0}_2$*, and hence, matrix* $\mathbf{M}$ *must be singular, which is stated as*

$$\det(\mathbf{M}) = 0$$

Upon expansion, the foregoing equation leads to

$$\det(\mathbf{M}) = 2x_1(-4x_1 + 3x_2) - 2x_2(9x_1 - 4x_2) = 0$$

or, after simplification,

$$x_1^2 + \frac{3}{2}x_1x_2 - x_2^2 = 0$$

thereby reducing the problem to the solution of two quadratic equations in two unknowns, the above equation and the third normality condition. Upon eliminating of x_2 *from the latter, and substitution of the expression thus resulting into the remaining equation, we obtain, after some simplifications,*

$$x_1^4 - x_1^2 + \frac{4}{25} = 0$$

which is, in fact, a quadratic equation in x_1^2, its roots being

$$(x_1^2)_{1,2} = \frac{1}{5}, \quad \frac{4}{5}$$

whence the four roots follow:

$$(x_1)_{1,2} = \pm\frac{\sqrt{5}}{5}, \quad (x_2)_{3,4} = \pm\frac{2\sqrt{5}}{5}$$

which yields, correspondingly,

$$(x_2)_{1,2} = \pm\frac{2\sqrt{5}}{5}, \quad (x_2)_{3,4} = \pm\frac{\sqrt{5}}{5}$$

More general problems of this family can be solved using the methods discussed in Ch. 4 for arbitrary objective functions subject to nonlinear equality constraints.

References

Angeles, J., 2002, *Fundamentals of Robotic Mechanical Systems. Theory, Methods, and Algorithms*, Second Edition, Springer-Verlag, New York.

Luenberger, D.G., 1984, *Linear and Nonlinear Programming*, Second Edition, Addison-Wesley Publishing Company, Reading (MA), pp. 345–359.

Salmon, G., 1964, *Higher Algebra*, reprinted from the 1885 Fifth Edition, Chelsea Publishing Co., New York.