The reduced gradient method for solving redundancy in robot arms

A. De Luca and G. Oriolo, Roma

Abstract. An efficient computational scheme for solving inverse kinematic problems in redundant robot arms is presented. When following a given end-effector trajectory, successive internal arm configurations are in general selected by local optimization of a given performance criterion. Typically, joint displacements are derived using the Projected Gradient method, involving pseudoinversion of the robot Jacobian and projection in its null-space. However, this technique is computationally intensive. In this paper, an alternative approach is proposed based on the Reduced Gradient method, which allows to deal explicitly only with the redundant degrees of freedom. The superiority of this technique for solving redundancy is illustrated analytically in a simple case, and numerically by simulation of a four-link planar arm. Optimization of various criteria like manipulability, available joint range, and distance from obstacles is considered. Extensions of the method are also briefly discussed.

Eine reduzierte Gradientenmethode zur Lösung des inversen kinematischen Problems redundanter Roboterarme


1 Introduction

The complexity of tasks that advanced manipulators are required to execute in a real-life environment asks for robot arms with increased, human-like dexterity. Robots should be able to move safely in a crowded workspace, with objects to be manipulated and obstacles to be avoided. Typical industrial examples of such situations include spray painting of the inside of car bodies and assembly tasks with multiple exchange of forces/torques with the environment.

To acquire dextrous capabilities, the arm structure should possess a number of degrees of freedom larger than the one strictly needed for the general positioning, i.e. should be redundant. In this case, the inversion of the direct kinematic relation

$$p = f(\mathbf{q})$$

between joint coordinates $\mathbf{q} \in \mathbb{R}^N$ and task coordinates $\mathbf{p} \in \mathbb{R}^M$, with $N > M$, yields an infinity of solutions. At a differential level, one has

$$d\mathbf{p} = J(\mathbf{q}) d\mathbf{q},$$

with the Jacobian $J$ being a non-square matrix.

In order to make the redundancy of the mechanical design really appealing, the robot controller should be programmed so to exploit as much as possible these extra degrees of freedom. The choice among the infinite number of joint trajectories realizing a given path of the end-effector can be done by optimizing a suitable criterion. Indeed, it is desirable to achieve a skilled robot performance with the minimum amount of additional complexity.

Kinematic redundancy is mainly used to enhance the functional workspace of the robot. In redundant arms, the crossing of kinematic singularities can be avoided through internal reconfiguration of the arm along the same given end-effector motion. This may be obtained by maximizing a criterion which measures the arm manipulability [1, 2]. Optimization schemes are useful also when obstacles are present in the robot workspace. While the end-effector trajectory is already planned so to avoid collision, redundancy can be used to guarantee a collision-free path for the whole arm, by means of various distance criteria [3–5]. There are also forbidden regions directly specified in the joint space: optimal avoidance of joint range limits has been considered by Liégeois (1977).

All the above techniques are based on the pseudoinverse of the Jacobian matrix, used to satisfy the end-effector dis-
placement, combined with the specification of a null-space vector according to the chosen objective. Liégeois [6] recognized this scheme to be the transposition of the Projected Gradient (PG) method to the robot inverse kinematic problem. The major flaw of this optimization algorithm is that it requires a considerable amount of computations. This limit was pointed out by several authors who provided various improvements. In particular, Klein and Huang [7] suggested the use of a Gaussian elimination technique within this approach; Khalil and Chevallereau [8] proposed a different way for computing the pseudoinverse, while Dubey et al. [9] implemented a faster numerical scheme for a seven-dof robot with a spherical wrist.

An alternative and more efficient approach is presented in this paper, based on the Reduced Gradient (RG) method for nonlinear constrained optimization. The resulting scheme solves redundancy in a natural way. In fact, only the extra degrees of freedom are used for optimization, while the remaining joint variables are in charge of the task satisfaction. This allows for large computational savings. In the following, the redundancy resolution problem will be formally revisited as an optimization problem. Comparison between the RG and PG methods will be carried out both analytically and numerically, showing the inherent benefits of the proposed approach.

2 Optimal redundancy resolution

For a redundant robot arm the optimal choice of joint configurations can be defined by the following nonlinear optimization problem:

\[
\begin{align*}
\max_q H(q), \quad \text{s.t.} \quad f(q) - p &= 0.
\end{align*}
\]

(3)

\(H(q)\) is a generic objective function, while the vector constraint in (3) follows from the direct kinematics of the arm. Typically, this optimization problem cannot be solved in a closed-form and an iterative process has to be devised. Starting from an initial point \(q^0\) in the joint space, an update is defined as:

\[
q^{k+1} = q^k + dq^k.
\]

(4)

If \(q^0\) is the actual arm configuration at some initial instant, the iterates (4) generated by the chosen optimization algorithm can be taken as reference values to be tracked by a closed-loop robot controller.

When \(p\) is constant, the solution (3) is the optimal posture \(q^*\) for the specified end-effector location. The optimization process will produce a sequence of intermediate joint configurations providing a self-motion of the arm. This is true as long as the constraint is never violated.

If a time-varying \(p\) is specified, i.e. a trajectory for the end-effector, (3) is transformed into a succession of subproblems obtained by discretization of the end-effector path into a sequence \(\{p^k\}\). Letting \(dp^k = p^{k+1} - p^k\), the \(k\)-th subproblem becomes of the form:

\[
\begin{align*}
\max_{dq} H(q^k + dq), \quad \text{s.t.} \quad f(q^k + dq) - (p^k + dp^k) &= 0.
\end{align*}
\]

(5)

Any optimization algorithm can be used for solving (5). In general, some iterations will be needed to find an exact solution to each subproblem, so that use of this scheme in real-time is possible only if the algorithm is extremely fast. Therefore, \(dq^k\) in (4) is usually chosen as the first iterate computed toward the optimum of (5).

One way of determining \(dq^k\) is given by the Projected Gradient method [10], namely by projecting the gradient of \(H(q)\) onto the tangent space of the constraint. Since \(\mathbb{R}^N = \mathcal{N}(J(J) \ominus \mathcal{H}(J^T))\) holds, the displacement \(dq^k\) is obtained as the sum of two vectors belonging to these complementary subspaces. The projection matrix in the null-space of the Jacobian is \(P = I - J^T J\), where \(J^T\) is the unique pseudoinverse of \(J\) (Boullion and Odell, 1971). Hence, the PG update results in

\[
dq^k = J^T dp^k + (I - J^T J) V_q H(q^k),
\]

(6)

where \(V_q H = (\partial H/\partial q)^T\) and all matrices are evaluated at \(q^k\).

The computations involved in (6) are quite cumbersome for redundant robots with several degrees of freedom. In fact, the pseudoinverse of \(J\) is obtained in general via a singular value decomposition technique. When the Jacobian has full row rank, \(J^T\) can be evaluated more directly as \(J^T(JJ^T)^{-1}\). However, this still requires a product of matrices and a \(M \times M\) matrix inversion. Moreover, the \(N \times N\) projection matrix \(P\) has rank equal to \((N - M)\) – the rank deficiency being a considerable waste of information.

3 The reduced gradient method

An alternative method for solving (5) is based on the observation that the actual number of free variables in the problem is only \(N - M\). Therefore, the search for an optimal displacement \(dq^k\) can be more efficiently performed within a reduced space of joint variables. This leads to the Reduced Gradient method, for which a preliminary assumption is needed.

Nondegeneracy Assumption. At every specified \(p\), there exists a \(q \in \mathbb{R}^N\) such that \(f(q) = p\), for which a partition can be found

\[
q = (q_a, q_b), \quad q_a \in \mathbb{R}^M, \quad q_b \in \mathbb{R}^{N-M},
\]

yielding a nonsingular matrix

\[
V_{q_a} f(q) = J_a^T(q).
\]

This assumption restricts the admissible end-effector poses, discarding those for which the Jacobian necessarily loses full row rank. According to the partition of \(q\), the Jacobian matrix is decomposed into two blocks \((J_a, J_b)\), so that the differential relation (2) can be rewritten as

\[
dp = J_a(q) dq_a + J_b(q) dq_b.
\]

(7)
The Implicit Function Theorem is invoked to locally explicit the constraint $f(q_a, q_b) - p = 0$. This provides an expression for the basic joint variables $q_a$ in terms of the independent variables $q_b$ and of the known end-effector pose $p$:

$$q_a = g(q_b, p).$$  \hfill (8)

From the same Theorem, an expression is found for the derivative of (8) with respect to $q_b$:

$$\frac{\partial g}{\partial q_b} = -J_{a}^{-1} J_{b}.$$  \hfill (9)

The reduced gradient of the objective function $H(q)$ with respect to $q_b$ is defined as the gradient of the function $H(q_b) = H(g(q_b, p), q_b)$, considered for a fixed $p$. From the chain rule

$$V_{q_b} H'(q_b) = \frac{\partial H}{\partial q_b} \bigg|_{q_a(g(q_b, p))} = \left[- (J_{a}^{-1} J_{b})^{T} \right] V_{q} H.$$  \hfill (10)

At iteration $k$, for the independent joint variables $q_b$, a step is taken in the reduced gradient direction:

$$d_{q_b}^{k} = V_{q_b} H'(q_b^{k}).$$  \hfill (11)

The displacement for the basic joint variables $q_a$ is then chosen so to satisfy (7):

$$d_{q_a}^{k} = J_{a}^{-1} dp - (J_{a}^{-1} J_{b}) \, d_{q_b}^{k}.$$  \hfill (12)

The overall update of the RG method can be written in a compact form as:

$$\begin{bmatrix} d_{q_a}^{k} \\ d_{q_b}^{k} \end{bmatrix} = \begin{bmatrix} J_{a}^{-1} \\ 0 \end{bmatrix} dp^{k} + \left[ \begin{bmatrix} J_{R} & J_{R}^{T} \\ -J_{R} & 1 \end{bmatrix} \right] V_{q} H(q^{k}),$$  \hfill (13)

where $J_{R} \triangleq J_{a}^{-1} J_{b}$. This should be compared for complexity with the update (6) of the PG method. It is worth noting that neither the pseudoinverse $J^\dagger$ nor the projection matrix $P$ are needed in (13). As a matter of fact, only one key matrix, i.e. $J_{R}$, has to be determined. The computational burden is essentially limited to the inversion of the $M \times M$ matrix $J_{a}$, which is directly available as part of the robot Jacobian.

It should be stressed that the PG and the RG methods do not provide in general the same updates in the overall space of joint configurations. The asymptotic rate of convergence is similar for the two methods, although for most problems the RG method will converge faster [10]. Moreover, each iteration of the RG is less time-consuming, so that its use in on-line implementations is more convenient.

The basic scheme outlined above can be further improved. In fact, for a given set of basic variables $q_a$, the inversion of the matrix $J_a$ in (12) may become ill-conditioned, thus affecting the convergence rate of the algorithm. To overcome this problem, the whole set of joint variables can be explored at each iteration, and the basic variables $q_a$ are chosen so to give the best conditioned $J_a$ matrix, e.g. using the absolute value of the determinant as a measure.

Another limiting factor, which is common to all methods based on the local linearization of the constraints, is that the generated update is feasible only to the first order. Uncompensated numerical errors build up, resulting in a deviation from the desired end-effector path. The inclusion of a correction procedure considerably improves constraint satisfaction. Denote by $q^{k+1}$ the temporary update provided by (13) in (4). An additional move $d_{q}^{k} = (\tilde{d}_{q_a}^{k}, 0)$ is performed in the direction orthogonal to the tangent plane at $q^{k}$,

$$\tilde{d}_{q_a}^{k} = -J_{a}^{-1} (q^{k}) \, \epsilon^{k+1},$$  \hfill (14)

using the 'expected' error $\epsilon^{k+1} = f(q^{k+1}) - p^{k+1}$. In principle, this step should be repeated several times to restore feasibility, especially when the constraint curvature is large. However, very good results have been obtained in numerical tests using a single correction step. At iteration $k$, the complete update for $q_a$ becomes

$$d_{q_a}^{k} = J_{a}^{-1} (dp - \epsilon^{k+1}) - (J_{a}^{-1} J_{b}) \, d_{q_b}^{k},$$  \hfill (15)

and $q^{k+1} = q^{k+1} + d_{q}^{k}$. This correction procedure makes the method inherently more robust, in much the same way as feeding back the cartesian error does in a closed-loop control scheme.

Finally, note that in (11) a unit stepsize was assumed along the reduced gradient direction. To improve convergence properties, a stepsize $\alpha \neq 1$ could be taken, either constant or generated by a line search.

4 Analytic results

In this section, a simple redundant robot will be used to illustrate how the RG method for solving redundancy leads to an efficient computational scheme. Moreover, it will be shown that the RG and PG methods generate the same search directions when used for optimizing a self-motion in robots with only one degree of redundancy.

Example 1 (3-R Robot). Consider a planar robot arm with three rotational joints, for which the degree of redundancy of the end-effector positioning task is one. For simplicity, all links are assumed of equal length $l$. Using as joint coordinates the absolute angles $q_i$ referred to the $x$-axis, the direct kinematics becomes

$$p_x = l(c_1 + c_2 + c_3), \quad p_y = l(s_1 + s_2 + s_3),$$  \hfill (16)

where the shorthand notation $s_i := \sin q_i, c_i := \cos q_i$ is used. The Jacobian matrix is

$$J(q) = \begin{bmatrix} -s_1 & -s_2 & -s_3 \\ c_1 & c_2 & c_3 \end{bmatrix}.$$  \hfill (17)

Note that a minor of this matrix, composed of columns $i$ and $j$, is nonsingular if $s_{i-j} := \sin (q_i - q_j)$ is nonzero.

Assuming full rank for the Jacobian, the pseudoinverse $J^\dagger$ and the null-space projection matrix $P$ are computed as:

$$J^\dagger = \frac{1}{l} \frac{1}{A} \begin{bmatrix} c_2 s_{2-1} + c_3 s_{3-1} & s_2 s_{2-1} + s_3 s_{3-1} \\ c_1 s_{1-2} + c_3 s_{3-2} & s_1 s_{1-2} + s_3 s_{3-2} \\ c_1 s_{1-3} + c_2 s_{2-3} & s_1 s_{1-3} + s_2 s_{2-3} \end{bmatrix}.$$
\[
P = \frac{1}{\Delta} \begin{bmatrix}
  s_{3-2} & s_{3-1} & s_{2-3} & s_{2-1} & s_{3-2} \\
  s_{3-1} & s_{3-2} & s_{3-1} & s_{2-1} & s_{1-3} \\
  s_{2-3} & s_{2-1} & s_{2-1} & s_{2-1} & s_{2-1} \\
  s_{1-3} & s_{2-1} & s_{2-1} & s_{2-1} & s_{2-1} \\
  s_{3-2} & s_{3-1} & s_{2-3} & s_{2-1} & s_{3-2}
\end{bmatrix}
\]

where \( \Delta = s_{2-1} + s_{3-1} + s_{2-2} \). These two matrices, used in (6), fully define the PG method.

To apply the RG method, any partition of the joint vector \( \mathbf{q} \) yielding a nonsingular matrix \( \mathbf{J}_a \) can be chosen. For instance, assuming \( s_{2-1} \neq 0 \), let \( \mathbf{q}_a = (q_1, q_2) \) and \( \mathbf{q}_b = q_3 \), which results in the partition of \( \mathbf{J} \)

\[
\mathbf{J}_a = \begin{bmatrix}
  -s_1 & -s_2 \\
  c_1 & c_2
\end{bmatrix}, \quad \mathbf{J}_b = \begin{bmatrix}
  -s_3 \\
  c_3
\end{bmatrix}.
\]  

(18)

Therefore, the matrix characterizing the RG method is

\[
\mathbf{J}_R = \frac{1}{s_{2-1}} \begin{bmatrix}
  s_{3-2} & s_{3-1}
\end{bmatrix}.
\]  

(19)

The reader may appreciate the simplicity of this derivation, which confirms the improvement obtained with the RG method. In this case, a step is computed as

\[
d\mathbf{q}^b = \left[ \mathbf{J}_a^{-1} \right] \mathbf{d} \mathbf{p}^b + \frac{1}{s_{2-1}} \begin{bmatrix}
  -s_{2-3} \\
  -s_{3-1} \\
  s_{2-1}
\end{bmatrix} \begin{bmatrix}
  -s_2 \\
  c_2 \\
  s_{2-1}
\end{bmatrix} V_\mathbf{q} H(\mathbf{q}^b).
\]

Indeed, similar expressions are obtained (modulo an index permutation) when a different set of basic variables is chosen.

**Example 2** (Self-motions of robots with \( N - M = 1 \)). For the class of robot arms with one degree of redundancy, under the nondegeneracy assumption, the Jacobian can always be partitioned as \( \mathbf{J} = (\mathbf{J}_a, \mathbf{j}_b) \), where \( \mathbf{J}_a \) is an \((N-1) \times (N-1)\) nonsingular matrix and \( \mathbf{j}_b \) is just a column. Suppose that a self-motion is performed to extremize an assigned criterion \( H(\mathbf{q}) \), at a given end-effector location \( \mathbf{p} \). In this case \( d\mathbf{p} = 0 \), and the only relevant matrix in the PG method is the matrix \( \mathbf{P} = \mathbf{I} - \mathbf{J}_R^{\top} \mathbf{J}. \) The following expression holds for the pseudo-inverse of a matrix partitioned as \( \mathbf{J} \) (Boillond and Odell, 1971)

\[
\mathbf{J}^p = \left[ \begin{bmatrix}
  \mathbf{J}_a^{-1} (1 - \beta \mathbf{j}_b \mathbf{j}_b^\top (\mathbf{J}_a \mathbf{J}_a^\top)^{-1}) \\
  \beta \mathbf{j}_b \mathbf{j}_b^\top (\mathbf{J}_a \mathbf{J}_a^\top)^{-1}
\end{bmatrix} \right],
\]  

(20)

where \( \beta \triangleq 1/(1 + \mathbf{j}_b^\top (\mathbf{J}_a \mathbf{J}_a^\top)^{-1} \mathbf{j}_b) \). After some manipulation, it can be shown that the projection matrix takes in this case the form

\[
\mathbf{P} = \beta \begin{bmatrix}
  (\mathbf{J}_a^{-1} \mathbf{j}_a)(\mathbf{J}_a^{-1} \mathbf{j}_a)^\top & -\mathbf{J}_a^{-1} \mathbf{j}_b \\
  -(\mathbf{J}_a^{-1} \mathbf{j}_a)^\top & 1
\end{bmatrix}.
\]  

(21)

Comparing (21) with the expression (13) of the RG method, as applied to self-motions, it is easy to see that

\[
d\mathbf{q}_{PG} = \beta \cdot d\mathbf{q}_{RG}.
\]  

(22)

Therefore, the two methods provide the same direction in the joint space, up to the iteration-dependent scaling factor \( \beta \). This was not unexpected, since for these robots the null-space of the Jacobian is one-dimensional. It is possible to show that this scaling factor is equivalently rewritten as

\[
\beta = \frac{\det^2 \mathbf{J}_a}{\det (\mathbf{J} \mathbf{J}^\top)} \leq 1.
\]  

(23)

This general result applies indeed to the previous example.

### 5 Numerical results

The proposed redundancy resolution method has been used for a 4-R planar robot arm with all links of unit length (Fig. 1), executing both internal self-motions and end-effector trajectories. For positioning tasks, this robot has two degrees of redundancy and thus it is suitable to fully illustrate the different behavior of the RG and the PG methods. In terms of absolute joint coordinates, the Jacobian of this arm is

\[
\mathbf{J}(\mathbf{q}) = \begin{bmatrix}
  -s_1 & -s_2 & -s_3 & -s_4 \\
  c_1 & c_2 & c_3 & c_4
\end{bmatrix}.
\]  

(24)

Different objective functions \( H(\mathbf{q}) \) were tested for optimization. In order to maximize the joint range availability, the criterion proposed by Liégeois [6] was used. Since relative joint angles are \( \theta_i = q_{i+1} - q_i \), for the 4-R arm this will be written as

\[
H_1(\mathbf{q}) = \frac{1}{2} \sum_{i=1}^{4} \left( q_{i+1} - \theta_i - \bar{\theta}_i \right)^2,
\]  

(25)

where \( [\theta_{m,i}, \theta_{M,i}] \) is the admissible range for joint \( i \) and \( \bar{\theta}_i \) is its center. In the first term of the summation, \( q_0 = 0 \). For obstacle avoidance, the case of a small disc centered at \( (\bar{x}, 0) \) has been considered. In particular, \( \bar{x} \) is such that only interference with the fourth link can occur. Properly modifying the criterion proposed by Bailieu (1986), it was chosen to maximize

\[
H_2(\mathbf{q}) = \sum_{i=1}^{3} \left( s_4 - s_{4-i} \right).
\]  

(26)

Finally, a test with a manipulability measure was carried out. For the 4-R arm an index which is equivalent to the

Fig. 1. A four-link planar robot arm
In the first numerical example the end-effector path is given as a straight line connecting point (2.93, 1) to (2.5, 0.7). The cartesian path is discretized with 200 samples and should be executed without violating the joint limits, which are set at ±90° of relative angle between successive links. At the starting point, the arm configuration is \( q = (15°, -15°, 0°, 90°) \), with the fourth joint at its upper limit. Figure 2a–d show the evolution of the four relative joint angles when using the PG method, the RG method or simple pseudo-inversion (PS). For the first two, maximization of \( H_1 \) is performed. The PS solution, plotted for comparison, violates the fourth joint limit since it naturally produces the minimum norm displacement of the joints. Instead, both PG and RG methods remain feasible along the whole path, although the latter drives sooner the fourth joint away from its limit. Note that, in correspondence to changes of the set of basic variables, a sudden slope variation occurs in the RG joint path.

In the second simulation, the robot task is to follow a vertical line from (3.41, 1.41) to (3.41, -0.3), in the presence of a disc obstacle of radius \( r = 0.1 \), located at \( x = 3.1 \). The arm starts with \( q = (0°, 0°, 45°, 45°) \). The natural motion generated without any optimization, i.e. using the PS method, collides as expected with the obstacle (Fig. 3). On the contrary, Fig. 4 shows the safe arm movement generated by the RG method maximizing the criterion \( H_2 \). The amount of computation needed for this obstacle avoidance scheme can be estimated in approx. 12 trigonometric evaluations, 11 products, and 10 additions for each iteration.

In the last numerical example, the arm performs a self-motion to recover manipulability, while keeping the end-effector fixed at the origin. The almost singular initial configuration \( q = (-3°, 3°, 177°, -177°) \) is shown in Fig. 5, together with the configuration \( q = (-45°, 45°, 135°, -135°) \), where the index \( H_3 \) assumes its maximum value. Both RG and PG optimization techniques yield this same final solution. However, as indicated in Fig. 6, the RG method is faster, achieving convergence within half of the PG iterations. In this case, the set of basic variables is \( (q_1, q_2) \) and does not change during the whole motion.
All simulations were performed with a constant stepsize \( \alpha = 0.1 \). Numerical evidence has shown that, when the correction term (14) is included, both RG and PG methods are less sensitive to the actual stepsize value.

**6 Conclusions**

A new computational scheme for solving redundancy in robot arms has been presented, based on the Reduced Gradient optimization method. The application of this technique to robotics relies on the physical intuition of the problem. The free selection of basic variables provides an explicit command on the choice of a particular inverse solution. With this respect, the approach can be recasted in the framework of inverse kinematic functions, as defined by Wampler [11]. The RG method has been applied here for local optimization of position-dependent criteria at a kinematic level. However, the method may be used also in global resolution schemes, as well as for revisiting other kinematic approaches like the task-priority method of Nakamura et al. [5]. In fact, the joint-space decomposition idea does not need the use of pseudoinversion and projection even when redundancy is resolved through task-augmentation [12]. Due to its generality, the presented scheme can be extended also to acceleration/torque resolution of redundancy [13], taking into account the full robot dynamic model. Finally, it is worth to mention that the explicit inversion of matrix \( J_R \) – the main requirement of the method – can be avoided by the proper application of the closed-loop scheme for coordinate transformations introduced in [14].

**Acknowledgements**

This work was partially supported by the Consiglio Nazionale delle Ricerche, grant no. 89.00521.67 (Progetto Finalizzato Robotica).

**References**


Dr. A. De Luca
G. Oriolo
Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Eudossiana 18
184 Rome
Italy

Received December 6, 1990