Effects of Different Objective Functions in Inequality Constrained and Rank-Deficient Least-Squares Problems

Lutz Roese-Koerner and Wolf-Dieter Schuh

Abstract Rank-deficient estimation problems often occur in geodesy due to linear dependencies or underdetermined systems. Well-known examples are the adjustment of a free geodetic network or a finite element approximation with data gaps. If additional knowledge about the parameters is given in form of inequalities (e.g., non-negativity), a rank-deficient and inequality constrained adjustment problem has to be solved.

In Roese-Koerner and Schuh (2014) we proposed a framework for the rigorous computation of a general solution for rank-deficient and inequality constrained least-squares problems. If the constraints do not resolve the manifold of solutions, a second minimization is performed in the nullspace of the design matrix. This can be thought of as a kind of pseudoinverse, which takes the inequality constraints into account.

In this contribution, the proposed framework is reviewed and the effect of different objective functions in the nullspace optimization step is examined. This enables us to aim for special properties of the solution like sparsity (L^1 norm) or minimal maximal errors (L^∞ norm). In a case study our findings are applied to two applications: a simple bivariate example to gain insight into the behavior of the algorithm and an engineering problem with strict tolerances to show its potential for classic geodetic tasks.

Keywords Convex Optimization · Inequality constrained least-squares · Rank defect · L^1 norm · L^2 norm · L^∞ norm · Nullspace minimization

Lutz Roese-Koerner · Wolf-Dieter Schuh
Institute of Geodesy and Geoinformation, University of Bonn
Tel.: +49-228-737423, Fax: +49-228-736486
e-mail: {roese-koerner, schuh}@geod.uni-bonn.de

1 Introduction

Geodesists often encounter rank-deficient optimization problems. This can be either due to underdetermined equation systems (e.g., resulting from a second order design of a geodetic network with more weights to be estimated than entries in the criterion matrix) or due to external parameters, which cannot be estimated from the observations (e.g., a datum defect). These cases do not result in one unique but in a manifold of solutions. If the situation gets even more complicated and additional knowledge about the parameters in form of linear inequality constraints is given, no closed formulas exist. Inequalities for nonnegative quantities like run-time delay in GPS, SAR or VLBI or e.g., slope constraints in surface fitting are typical constraints.

In the unconstrained case, usually a second objective function is introduced to the problem to enforce a unique solution despite the rank-defect. A classic choice would be to minimize the length of the solution vector with respect to the L^2 norm (cf. Gill et al, 1991, p.230-234). While unconstrained problems with a rank-defect are well studied, little is known about the inequality constrained case.

The special case of non-negative least-squares with a possible rank-defect and additional inequalities was solved by Schaffrin (1981). However, this method cannot be generalized easily. Werner and Yapar (1996) proposed a method for computing a rigorous general solution of inequality constrained problems with a rank-defect. Unfortunately, their method is solely suited for small-scale problems as it involves arbitrarily testing of subsets of the constraints, which becomes a limiting factor in the multivariate case. Additional work by Xu et al (1999) is focused on the stabilization of ill-conditioned linear complementarity problems. Dantzig (1998) described a method to compute one arbitrary particular solution despite a possible rank-defect for problems with a linear or quadratic objective function. However, no description of the manifold of solutions is given.

All these algorithms are either tailor-made for special problems or yield only one of an infinite number of particular solutions or are restricted to small-scale problems. Therefore, in Roese-Koerner and Schuh (2014) we developed a framework for the rigorous computation of a general solution of inequality constraint least-squares problems. Based on this framework, here we set out to compute a solution with desirable properties like e.g., sparsity, exploring the opportunities which result from the choice of the objective function in the second minimization step.

2 Inequality constrained estimation

In the following, we focus on a linear Gauss-Markov model (GMM)

$$\boldsymbol{\ell} + \mathbf{v} = \mathbf{A}\mathbf{x}. \tag{1}$$

$n \times 1$ vector $\boldsymbol{\ell}$ contains the observations and $n \times 1$ vector $\mathbf{v}$ the corresponding residuals. $\mathbf{A}$ is the design matrix and $\mathbf{x}$ comprises the m parameters to be estimated. Following the least-squares principle, the (weighted) sum of squared residuals shall be minimized

$$\Phi(\mathbf{v}) = \mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{v} \ldots \min. \tag{2}$$

$\boldsymbol{\Sigma}$ is the possibly fully populated variance-covariance matrix of the observations. The objective function shall be minimized with respect to p inequality constraints which have to be fulfilled strictly. Linear inequality constraints can be formulated as

$$\mathbf{B}^T \mathbf{x} \leq \mathbf{b}. \tag{3}$$

$m \times p$ matrix $\mathbf{B}$ is called constraint matrix and $p \times 1$ vector $\mathbf{b}$ is the corresponding right-hand-side of the constraints. The constraints can be subdivided in active constraints $\mathbf{B}_a, \mathbf{b}_a$, which hold as equality constraints at the optimal solution $\widetilde{\mathbf{x}}$ and inactive constraints $\mathbf{B}_i, \mathbf{b}_i$, which hold as strict inequalities

$$\mathbf{B}_a^T \mathbf{x} = \mathbf{b}_a, \qquad \mathbf{B}_i^T \mathbf{x} < \mathbf{b}_i. \tag{4}$$

Minimizing (2) subject to (3) will be refered to as the inequality constrained least-squares (ICLS) problem, which can be expressed as a quadratic program (QP) in standard form

INEQUALITY CONSTRAINED LEAST-SQUARES

$$\begin{aligned} &\textbf{objective funct.:}\ \Phi(\mathbf{x}) = \mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{v} \ldots \text{Min} \\ &\textbf{constraints:}\quad \mathbf{B}^T \mathbf{x} \leq \mathbf{b} \\ &\textbf{optim. variable:}\ \mathbf{x} \in \mathbb{R}^m. \end{aligned} \tag{5}$$

A QP is a convex optimization problem with a quadratic objective function and linear constraints. As QPs are well studied, there is a variety of existing algorithms to solve them efficiently. Therefore, it can be beneficial to reformulate the described problem as a QP. As it is not known beforehand which constraints will be active in the optimal solution, only iterative algorithms exist. Most of them can be subdivided into two classes: simplex methods (e.g Dantzig's simplex method for quadratic programming, Dantzig, 1998, p. 490-498) and interior-point methods (e.g., primal-dual methods, Boyd and Vandenberghe, 2004, p. 609-613).

3 Rank deficient ICLS problems

In the following, we focus on determining a unique rigorous general solution of the ICLS problem (5) with a rank-deficient design matrix $\mathbf{A}$,

$$Rg(\mathbf{A}) = r < m, \qquad d = m - r. \tag{6}$$

The proposed framework consists of three major parts. First, inequality constraints are not taken into account, a general solution of the rank-deficient unconstrained problem is computed, and a transformation of parameters is performed (described in Sect. 3.1). The next step depends on whether the manifold of solutions and the feasible set intersect. In case of an intersection (*case 1*, Sect. 3.2), there is still a manifold and a second minimization procedure is carried out in the nullspace. If there is no such intersection (*case 2*, Sect. 3.3), Dantzig's method is used to compute a particular solution of the original problem. The resulting particular solution can be shown to be unique if no active constraint is parallel to the manifold. In turn, an active parallel constraint calls for a nullspace optimization.

3.1 Transformation of parameters

As described in Roese-Koerner and Schuh (2014), the introduction of linear inequality constraints can result in a shift and/or a restriction of the manifold but never in a rotation. Therefore it is instructive to first compute a general solution of the unconstrained ordinary least-squares (OLS) problem

$$\widetilde{\mathbf{x}}^{OLS}(\boldsymbol{\lambda}) = \mathbf{x}_P^{OLS} + \mathbf{X}_{hom}\,\boldsymbol{\lambda}. \tag{7}$$

Subsequently, the constraints (3) can be reformulated with respect to the d free parameters $\boldsymbol{\lambda}$, the particular solution $\mathbf{x}_P^{OLS}$ and the homogenous solution $\mathbf{X}_{hom}$,

$$\mathbf{B}^T\left(\mathbf{x}_P^{OLS} + \mathbf{X}_{hom}\,\boldsymbol{\lambda}\right) \leq \mathbf{b} \tag{8}$$

With the substitutions $\mathbf{B}_{\boldsymbol{\lambda}}^T := \mathbf{B}^T\mathbf{X}_{hom}$ and $\mathbf{b}_{\boldsymbol{\lambda}} := \mathbf{b} - \mathbf{B}^T\mathbf{x}_P^{OLS}$, (3.1) reads

$$\mathbf{B}_{\boldsymbol{\lambda}}^T\boldsymbol{\lambda} \leq \mathbf{b}_{\boldsymbol{\lambda}}. \tag{9}$$

If these constraints are contradictory (as it can be examined by solving a feasibility problem, cf. Boyd and Vandenberghe, 2004, p.579-580), there is no intersection of manifold and feasible set and we proceed as described in Sect. 3.3. Otherwise, we proceed as described in Sect. 3.2.

3.2 Case 1: Intersection

In case of an intersection, we aim for the rigorous computation of a unique particular solution. Therefore, a second optimization problem is introduced (e.g., minimizing the length of the solution vector with respect to the norm L^p). As this minimization takes place in the nullspace of design matrix $\mathbf{A}$, the value of the objective function (2) of the original problem does not change

$$\boxed{\begin{array}{ll} \text{NULLSPACE OPTIMIZATION PROBLEM} & \\ \textbf{objective funct.:} & \Phi_{NS}\left(\mathbf{x}_P^{ICLS}(\boldsymbol{\lambda})\right) \ldots \text{Min} \\ \textbf{constraints:} & \mathbf{B}_{\boldsymbol{\lambda}}^T\boldsymbol{\lambda} \leq \mathbf{b}_{\boldsymbol{\lambda}} \\ \textbf{optim. variable:} & \boldsymbol{\lambda} \in \mathbb{R}^d. \end{array}}. \tag{10}$$

The minimization yields optimal free parameters $\widetilde{\boldsymbol{\lambda}}$, whose insertion in (7) results in

$$\widetilde{\mathbf{x}}_P^{ICLS} = \mathbf{x}_P^{OLS} + \mathbf{X}_{hom}\widetilde{\boldsymbol{\lambda}}, \tag{11}$$

which is a unique particular solution that fulfills all constraints.

3.3 Case 2: No intersection

If manifold and feasible region are disjunct and there is no active parallel constraint, the constraints have resolved the manifold. Therefore, it is sufficient to compute one particular solution of the constraint problem — e.g., with Dantzig simplex method for QPs. It can be shown that

$$\widetilde{\mathbf{x}}^{ICLS} = \mathbf{x}_P^{ICLS} \tag{12}$$

is the unique solution. Instead, if at least one active constraint is parallel to the manifold, there is a shift of the manifold and a second objective function has to be introduced, as described in Sect. 3.2.

A more detailed description of the framework is provided in Roese-Koerner and Schuh (2014).

4 Nullspace optimization

In this section, the choice of the (second) objective function Φ_{NS} in the nullspace optimization problem (10) is discussed. As the whole minimization takes place in the nullspace of the design matrix, the value of the original objective function (2) will not change. Nonetheless, choosing a suitable second objective function for a particular problem can be helpful to achieve properties like sparsity of the parameter vector.

Two different parts of the objective function can be distinguished, which will be examined in some detail: the *functional relationship* and the *norm*, with respect to which the minimization (or maximization) shall be performed.

4.1 Functional relationship

The *functional relationship* strongly depends on the application. However, there are two main concepts which are applicable to a big variety of problems.

First, the length of the parameter vector $\mathbf{x}$ can be minimized. This can e.g., be beneficial if not absolute coordinates but coordinate differences are estimated, which should be close to the initial coordinates.
More sophisticated approaches include a weighted minimization of the length of the parameter vector. For example if prior knowledge about the magnitude of the parameters is given. Prominent examples are Kaula's rule of thumb in gravity field estimation or the demand for a decay of the amplitudes of higher frequencies in signal processing to achieve a square integrable function. A mathematical example for the minimization of the length of the parameter vector is provided in Sect. 5.1.

The second main concept is to maximize the distance to the constraints

$$||\mathbf{B}^T \mathbf{x}(\boldsymbol{\lambda}) - \mathbf{b}|| \dots \max \tag{13a}$$

$$\Longleftrightarrow ||\mathbf{B}^T \mathbf{x}_P + \mathbf{B}^T \mathbf{X}_{hom}\boldsymbol{\lambda} - \mathbf{b}|| \dots \max\,. \tag{13b}$$

This could be beneficial, if the constraints constitute a kind of outermost threshold. One of the main advantages of the use of inequality constraints is, that they do not influence the result, if they are not active. Therfore, it it usually not possible to provide a buffer to the boundary of the feasible set without losing estimation quality (shown by an increased value of the original objective function). However, due to the optimization in the nullspace of $\mathbf{A}$, we are in the unique position to apply such a buffer without deteriorating the estimate. In Sect. 5.2 an application is described, in which this type of *functional relationship* is applied.

4.2 Norms

In the following, we will point out some general aspects of different *norms* and their influence on the most classical *functional relationship*: the length of the solution vector. L^p norms

$$||\mathbf{x}||_p := \left(x_1^p + x_2^p + \ldots + x_m^p\right)^{1/p}, \quad p = 1, 2, \ldots, \infty$$

used in adjustment theory include the L^1 norm, the L^2 norm and the L^∞ norm (cf. Jäger et al, 2005; Boyd and Vandenberghe, 2004, p.125-128 and p. 635, respectively).

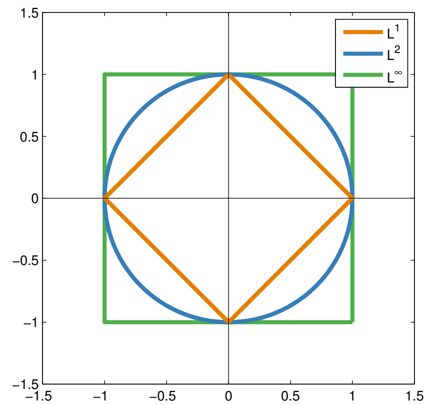


Fig. 1 Two dimensional *unit spheres* of L^1 (*orange*), L^2 (*blue*) and L^∞ norm (*green*). The L^1 diamond evolves from the summation of x and y value, while for the L^∞ square, only the value of the biggest quantity is decisive. Figure modified and extended from Tibshirani (1996).

Most often, the length of a vector is minimized with respect to the L^2 norm (also known as Euclidean norm). This is a quite natural choice, which can easily be visualized geometrically. However, as the influence of one element on the norm decreases if its absolute value becomes smaller, a minimization with respect to the L^2 norm will seldom result in a sparse vector. This can be verified by examining the *blue* L^2 norm unit sphere depicted in Fig. 1.

A naive choice to achieve maximal sparsity of a vector would be a minimization with respect to the L^0 norm (e.g., the number of nonzero elements). However, a minimization with respect to the L^0 norm is a combinatorical problem and computationally very demanding. Therefore, e.g., Candes et al (2006) approximated the L^0-minimization-problem by the L^1-minimization-problem in the context of compressed sensing. They showed, that a minimization with respect to the L^1 norm in most cases yields sparse results, too. This is used in many compressed sensing algorithms. Minimization with respect to the L^1 norm is a convex optimization problem and can be formulated as linear program (cf. Dantzig, 1998, p.60-62). As can be seen in Fig. 1, the corners of the *orange* L^1 norm unit sphere coincide with the coordinate grid. This is equivalent to the statement, that one parameter is zero there, yielding a sparse solution.

Application of the L^∞ norm (also called Chebyshev norm) results in a parameter vector with minimal maximal value, that is usually not sparse. See Fig. 1 for the corresponding unit sphere. The L^∞ norm is often applied, when trying to maximize the distance to the constraints (cf. Sect. 4.1), in order to maximize the minimal buffer to the boundary.

5 Case studies

The effect of a nullspace optimization will be demonstrated in two examples: a very simple bivariate one and a more sophisticated network adjustment problem.

5.1 Case Study 1: Bivariate Example

In this case, a least-squares estimate of the two summands of a weighted sum is to be calculated

$$\ell_i + v_i = x_1 + 2x_2 \tag{14}$$

subject to the constraints

$$x_1 \leq 2, \quad x_2 \leq 10 \tag{15}$$

using the framework described in Sect. 3. A contour plot of the objective function and the constraints is given in Fig. 2. *Red lines* represent constraints, the infeasible region is *shaded* and the *dashed black line* indicates the manifold of solutions. We assume the observations

$$\boldsymbol{\ell}^T = \begin{bmatrix} 23.2 & 16.4 & 12.9 & 8.2 & 13.7 \end{bmatrix}, \tag{16}$$

to be uncorrelated. The observation equations read

$$\boldsymbol{\ell} + \mathbf{v} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{A}\mathbf{x}, \tag{17}$$

with a clearly rank-deficient design matrix $\mathbf{A}$. Setting up the normal equations and applying the Gauss-Jordan algorithm yields a general OLS solution

$$\mathbf{x}(\lambda) = \begin{bmatrix} 14.88 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \end{bmatrix} \lambda = \mathbf{x}_P + \mathbf{X}_{hom}\lambda. \tag{18}$$

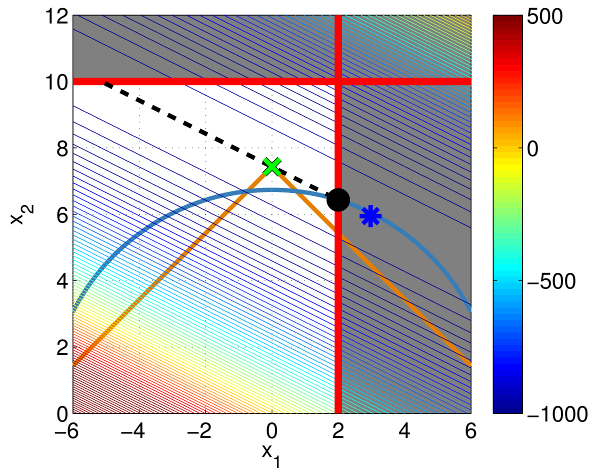


Fig. 2 Contour lines of the objective function of example 1. *Red lines* represent constraints, the infeasible region is *shaded*. The *dashed black line* indicates the manifold of solutions. The *green cross* is the L^1 norm solution, the *black circle* the L^2 solution. For comparison the solution using a pseudoinverse (*blue star*) is shown, too. Appropriately scaled unit spheres are depicted in *orange* (L^1) and *light blue* (L^2).

As there is an intersection of manifold and feasible set (cf. Fig. 2), the constraints are reformulated with respect to the free parameter λ and a second optimization problem in the nullspace of the design matrix has to be solved. We chose to examine the different effects of minimizing the length of the parameter vector with respect to the L^1 or L^2 norm

2D EXAMPLE: NULLSPACE OPTIMIZATION

objective funct.: $||\mathbf{x}_P + \mathbf{X}_{hom}\lambda||_p \dots \text{Min}$

constraints: $\mathbf{B}_\lambda^T \lambda = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \lambda \leq \begin{bmatrix} -12.88 \\ 10 \end{bmatrix} = \mathbf{b}_\lambda$

optim. variable: $\lambda \in \mathbb{R}$.

Depending on the chosen norm this results in either

$$\widetilde{\mathbf{x}}_P^{ICLS,L^2} = \begin{bmatrix} 2.00 \\ 6.44 \end{bmatrix} \quad \text{or} \quad \widetilde{\mathbf{x}}_P^{ICLS,L^1} = \begin{bmatrix} 0.00 \\ 7.43 \end{bmatrix}. \tag{19}$$

As expected, utilization of the L^1 norm in the nullspace optimization step yields in a sparse solution without an increase in the sum of squared residuals of the original problem.

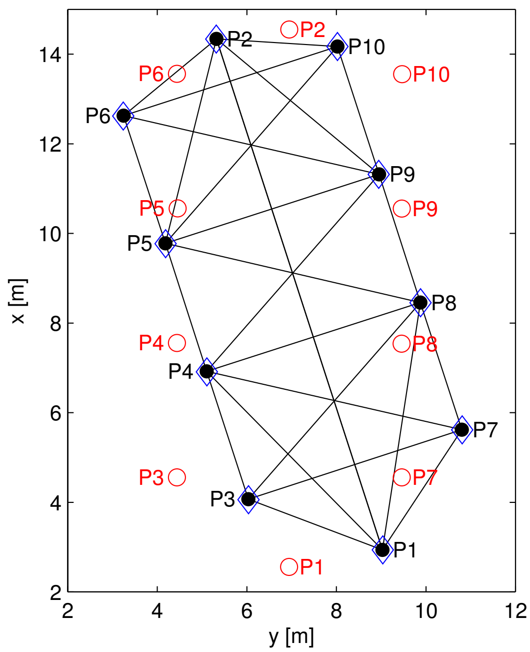


Fig. 3 Example 2: Distance measurements (*black lines*) are performed between points $P1$ to $P10$ (*black dots*). A particular OLS solution (*blue diamonds*) and the ICLS solution (*red circles*) with maximal minimal distance to the constraints are shown.

5.2 Case Study 2: Network Adjustment

The second case study is based on an engineering problem. We assume that some prefabricated building material shall be fitted between other elements so that the parts can be welded together. In order to make welding possible, tolerances have to be fulfilled strictly.

Fig. 3 depicts the test case. 26 distance measurements (*black lines*) are performed between the ten points $P1$ to $P10$ (*black dots*). Their 20 coordinates are the parameters to be estimated in a GMM (1). As no datum is defined, estimating absolute values of the coordinates is a rank-deficient problem.

Points $P3$ to $P6$ are located at the left-hand side of the gap the new part is supposed to fill, and the points $P7$ to $P10$ are located on its right-hand side. $P1$ and $P2$ are external points to stabilize the network. It shall be determined if the new part fits between both lines of points. This can be achieved by setting up the 16 linear constraints

$$y_{7,8,9,10} - y_{3,4,5,6} \leq 5.03m \tag{20}$$

and the 16 linear constraints

$$y_{7,8,9,10} - y_{3,4,5,6} \geq 5.00m, \tag{21}$$

resulting in a rank-deficient ICLS problem in form of (5). While the first constraints guarantee, that the new part is not allowed to be wider than 5.03m, the latter assure, that it is not smaller than 5.00m (otherwise the gap would be too big for welding). The constraints force the estimated points to align almost parallel to the x axis (cf. *red circles* in Fig. 3). If more than two of the 32 constraints mentioned above are active, the new part will not meet the tolerances. Incompatible elements can be detected via an analysis of the Lagrange multipliers (cf. Roese-Koerner et al, 2012).

If the manifold is not resolved through the introduction of constraints, a nullspace optimization has to be performed. This can be used to maximize the minimal distance to the constraints

$$\Phi_{NS,\infty} = ||\mathbf{B}^T \mathbf{x}(\boldsymbol{\lambda}) - \mathbf{b}||_\infty \ldots \max\,. \tag{22}$$

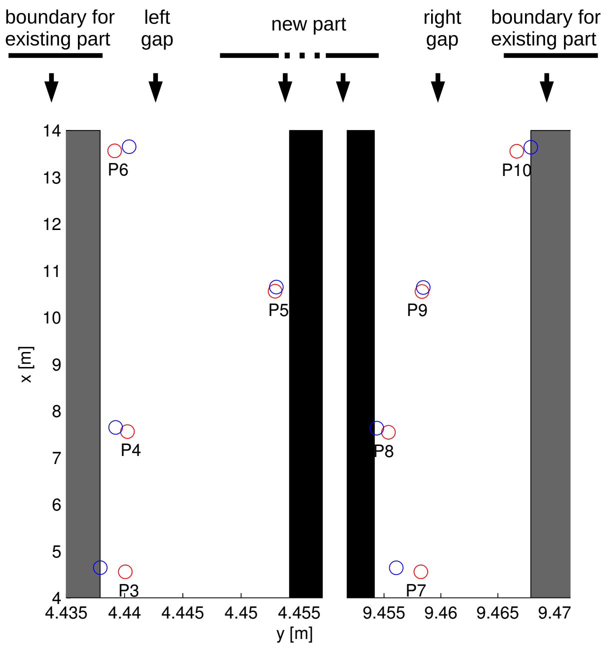
Using the Chebyshev norm is always beneficial if tolerances instead of standard deviations are given. The optimization problem was solved using the CVX software (Grant and Boyd, 2008). Results are shown in Fig. 3. In the chosen scenario, no constraint is active. Therefore, the new part will fit in the gap and welding is possible.

Fig. 4 shows the welding boundary for the existing parts (*gray area*), the new part (*black area*) and the "gaps" at its left-hand (Fig. 4(a)) and right-hand (Fig. 4(b)) side. Please note the different scales in x and y direction and the breach in the y axis. Adjusted coordinates of the ICLS estimate with $\Phi_{NS,\infty}$ (*red circles*) are compared with those of an ICLS estimate with

$$\Phi_{NS,L^2} = ||\mathbf{x}||_{L^2} \ldots \min \tag{23}$$

as nullspace objective function (*blue circles*).

While both estimates provide a decision *if* the new part will fit, only the adjustment with maximal minimal distance to the constraints allows to determine *how well* the new part will fit. This can be seen in Fig. 4, where for this estimate the minimal distance to the constraints is at least 2.5mm at each side (namely for the points 5,6,8 and 10). In contrast, the blue points 3 and 10 are exactly on the boundary. So there is clearly a benefit in choosing a suited objective function for the nullspace optimization.



(a) Gap on the left side. (b) Gap on the right side.

Fig. 4 Boundary for the existing parts (*gray area*), the new part (*black area*) and the "gaps" at the left-hand (a) and right-hand (b) side of the new part. The axes are scaled differently and there is a breach in the y axis, so most of the new part is not shown. Adjusted coordinates of the above mentioned ICLS estimate with maximal minimal distance to the constraints (*red circles*) are compared with those of an ICLS estimate with Φ_{NS,L^2} as nullspace objective function (*blue circles*).

Due to space limitations we restricted ourselves to a brief description of the application and the figurative results presented in Fig. 3 and Fig. 4. Further information (e.g., the observations and the functional and stochastic model) as well as quantitative results can be obtained from the authors.

6 Conclusion

A framework for the computation of a rigorous general solution of rank-deficient ICLS problems has been reviewed. It has been shown that it is possible to obtain a solution with certain predefined optimality properties, if the manifold of solutions is not resolved by the constraints. As this results from a minimization process in the nullspace of the design matrix, the sum of squared residuals remains unchanged. Therefore, the described approach can be beneficial for practical applications as useful properties, like e.g., sparsity or maximal minimal distances, can be obtained without sacrificing estimation quality.

References

Boyd S, Vandenberghe L (2004) Convex Optimization. Cambridge University Press

Candes E, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. Information Theory, IEEE Transactions on 52(2):489–509, DOI 10.1109/TIT.2005.862083

Dantzig G (1998) Linear Programming and Extensions. Princeton University Press, New Jersey, USA

Gill P, Murray W, Wright M (1991) Numerical Linear Algebra and Optimization, Vol. 1. Addison Wesley Publishing

Grant M, Boyd S (2008) Graph implementations for nonsmooth convex programs. In: Blondel V, Boyd S, Kimura H (eds) Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences, Springer-Verlag Limited, pp 95–110

Jäger RR, Müller T, Saler H, Schwäble R (2005) Klassische und robuste Ausgleichungsverfahren. Wichmann-Verlag, Heidelberg

Roese-Koerner L, Schuh WD (2014) Convex optimization under inequality constraints in rank-deficient systems. Journal of Geodesy, in press, DOI 10.1007/s00190-014-0692-1

Roese-Koerner L, Devaraju B, Sneeuw N, Schuh WD (2012) A stochastic framework for inequality constrained estimation. Journal of Geodesy 86(11):1005–1018, DOI 10.1007/s00190-012-0560-9

Schaffrin B (1981) Ausgleichung mit Bedingungs-Ungleichungen. Allgemeine Vermessungs-Nachrichten 88. Jg.:227–238

Tibshirani R(1996) Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society 58(1):267–288

Werner HJ, Yapar C (1996) On inequality constrained generalized least squares selections in the general possibly singular Gauss-Markov model: A projector theoretical approach. Linear Algebra and its Applications 237/238:359–393, DOI 10.1016/0024-3795(94)00357-2

Xu P, Cannon E, Lachapelle G (1999) Stabilizing ill-conditioned linear complementarity problems. Journal of Geodesy 73:204–213, DOI 10.1007/s001900050237