**FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG**
INSTITUT FÜR INFORMATIK (MATHEMATISCHE MASCHINEN UND DATENVERARBEITUNG)

**Lehrstuhl für Informatik 10 (Systemsimulation)**

# On a regularization technique for Kovarik-like approximate orthogonalization algorithms

Aurelian Nicola, Constantin Popa, Ulrich Rüde

Lehrstuhlbericht 09 - 14

# On a regularization technique for Kovarik-like approximate orthogonalization algorithms

Aurelian Nicola, Constantin Popa[*], Constantin Popa[†], Ulrich Rüde[‡]

**Abstract**

In this paper we consider four versions of Kovarik's iterative orthogonalization algorithm, for approximating the minimal norm solution of symmetric least squares problems. Although the theoretical convergence rate of these algorithms is at least linear, in practical applications we observed that a too large number of iterations can dramatically deteriorate the already obtained approximation. In this respect we analyze the above mentioned Kovarik-like methods according to the modifications they make on the "machine zero" eigenvalues of the problem's (symmetric) matrix. We establish a theoretical almost optimal formula for the number of iterations necessary to obtain an enough accurate approximation, as well as to avoid the above mentioned troubles. Experiments on collocation discretization of a Fredholm first kind integral equation illustrate the efficiency of our considerations.

**2000 MS Classification:** 65F10, 65F20

**Key words and phrases:** Kovarik-like algorithms, approximate orthogonalization, regularization, first kind integral equations.

## 1 Introduction

In the paper [2] Z. Kovarik proposed two iterative algorithms for approximate orthogonalization of a finite set of linearly independent vectors from a Hilbert space. In [5] and [6], we extended these algorithms for approximate orthogonalization of rows and columns of arbitrary rectangular matrices and in [3, 4, 7] we adapted these methods for symmetric matrices and for approximating the minimal norm solution $x_{LS}$ of symmetric least squares problems of the form

$$\| Ax - b \| = \min\{\| Az - b \|, z \in I\!R^n\} \tag{1}$$

with $A$: $n \times n$ symmetric and $\| \cdot \|$ the Euclidean norm. In this respect, the following four algorithms were obtained: let $A_0 = A, b^0 = b$; for $k = 0, 1, \ldots$ do until convergence

**Algorithm KOAS-rhs**
$H_k = I - A_k$
$A_{k+1} = (I + 0.5H_k)A_k$
$b^{k+1} = (I + 0.5H_k)b^k$

**Algorithm KOBS-rhs**
$K_k = (I - A_k)(I + A_k)^{-1}$
$A_{k+1} = (I + K_k)A_k$
$b^{k+1} = (I + K_k)b^k$

[*]"Ovidius" University of Constanta, Faculty of Mathematics and Computer Science, Romania; e-mail: anicola@univ-ovidius.ro.

[†]"Ovidius" University of Constanta, Faculty of Mathematics and Computer Science, Romania; e-mail: cpoopa@univ-ovidius.ro.; for this author the paper was supported by a DAAD Grant in the period 01.07 - 31.08.2009 at Institute for Computer Science 10 (Systemsimulation), Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

[‡]Institute for Computer Science 10 (Systemsimulation), Friedrich-Alexander-Universität Erlangen-Nürnberg, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; e-mail: ulrich.ruede@informatik.uni-erlangen.de

**Algorithm MKOBS-rhs**
$K_k = (I - A_k)(I - A_k + A_k^2)$
$A_{k+1} = (I + K_k)A_k$
$b^{k+1} = (I + K_k)b^k$

**Algorithm IFKOBS-rhs**
$K_k = (I - A_k)(I - 0.5A_k)$
$A_{k+1} = (I + K_k)A_k$
$b^{k+1} = (I + K_k)b^k$

**Theorem 1** *Let us suppose that $A$ is also positive semidefinite and its spectral radius $\rho(A)$ satisfies*

$$\rho(A) < 1. \tag{2}$$

*Then, any of the above algorithms generates sequences with the following properties:*
*(i) If the problem* (1) *is consistent, then* $\lim_{k\to\infty} b^k = x_{LS}$*;*
*(ii) If the problem* (1) *is inconsistent, then* $\lim_{k\to\infty} A_k b^k = x_{LS}$*. Moreover, in this case we have* $\lim_{k\to\infty} \parallel b^k \parallel = \infty$*.*

The previous algorithms have linear or superlinear convergence and a mesh independent behavior for problems arising from collocation discretization of Fredholm first kind integral equations (see [1, 3, 4]). Moreover, the assumption (2) is not restrictive because it can be obtained e.g. by the scaling

$$A = \frac{1}{1+ \parallel A \parallel_\infty} A. \tag{3}$$

But, in spite of these good properties, we have observed in practical computations that a too large number of iterations can deteriorate the computed solution dramatically. This "divergent behavior" can be observed very well in the following example. Let us consider the integral equation $\int_0^1 k(s,t)\, x(t)dt = y(s),\ s \in [0,1]$, discretized by the collocation method from [1] with the collocation points $s_i = \frac{i-1}{n-1}, i = 1, \ldots, n$. Then, the associated least squares formulation (1) has the elements $A$ and $b$ given by: $(A)_{ij} = \int_0^1 k(s_i,t)k(s_j,t)dt$ and $(b)_i = y(s_i)$. We used the kernel $k(s,t) = \frac{1}{1+|s-0.5|+t}$, for which the matrix coefficients $(A)_{ij}$ can be analytically obtained and computed. The $n \times n$ matrix $A$ is symmetric and positive semidefinite, with $rank(A) = \frac{n}{2}$ ($n$ even) or $\frac{n+1}{2}$ ($n$ odd) (see e.g. [3, 4]). If the right hand side $y(s)$ is such that $x_{ex}(t) = 1, \forall t \in [0,1]$ is a solution of the initial integral equation, we obtain a consistent least squares formulation as (1), which we shall denote by **P-cons**. Then, by keeping the matrix $A$ unchanged and adding a perturbation to the right hand side $b$ (in Matlab notation) $rand('state', 0); pert = rand(n, 1); b = b + pert$; we get an inconsistent version of (1), denoted by **P-pert**. We applied the above four algorithms to both problems **P-cons** and **P-pert**, by using different numbers of iterations and for fixed $n = 32$. Below we present the behavior of the residuals ($log(\parallel Ab^k - b \parallel)$ for **P-cons** in figures 1 - 2, and $log(\parallel A(A(Ab^k) - b) \parallel)$ for **P-pert** in figures 3 - 4) versus the number of iterations. We can observe that in each case there exists a "critical value", such that if the number of iterations exceeds it, the norm of the residual begins to increase and the computed approximation deteriorates dramatically (the values of this critical value for KOAS-rhs, KOBS-rhs, MKOBS-rhs, IFKOBS-rhs are 88, 51, 51, 52 for **P-cons** and 48, 29, 32, 32 for **P-pert**, respectively).

## 2 The analysis of the divergent behavior

In order to understand and analyze the unpleasant behavior mentioned in section 1 we first consider the following general description of the above four algorithms.
**Algorithm** $\mathcal{K}$-rhs. Let $A_0 = A, b^0 = b$; for $k = 0, 1, \ldots$ do until convergence

$$A_{k+1} = f(A_k)A_k, \quad b^{k+1} = f(A_k)b^k, \tag{4}$$

where the function $f$ is given by

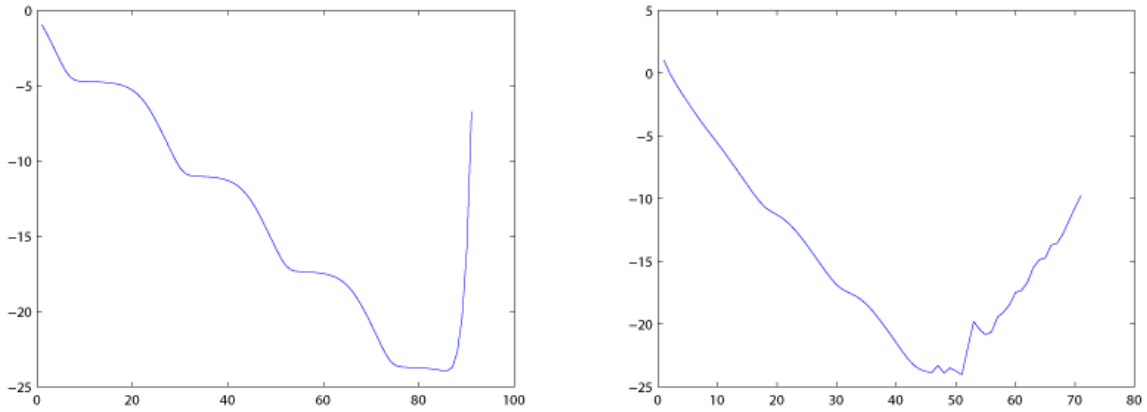


Figure 1: **P-cons; n = 32**: KOAS-rhs (left) and KOBS-rhs (right)

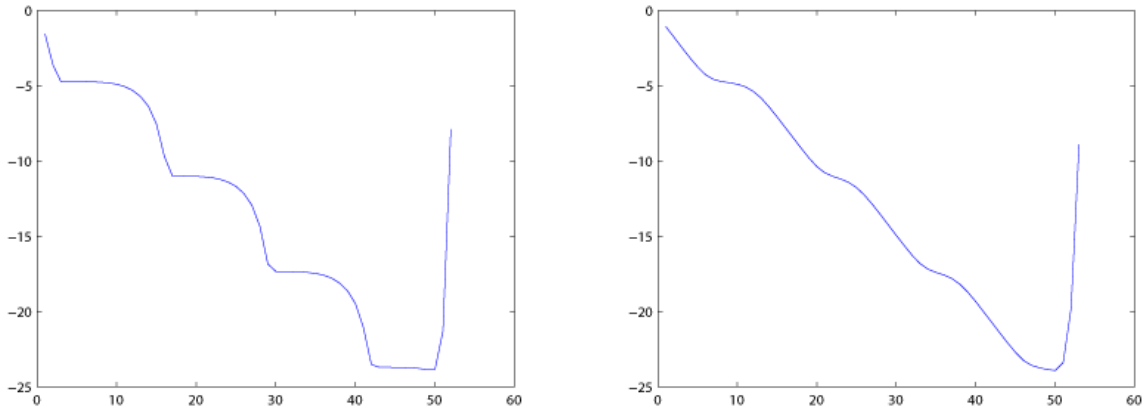


Figure 2: **P-cons; n = 32**: MKOBS-rhs (left) and IFKOBS-rhs (right)

$$f(x) \;=\; \begin{cases} \frac{2}{1+x}, & \text{if } \; \mathcal{K} = KOBS \\ 1+\frac{1}{2}(1-x), & \text{if } \; \mathcal{K} = KOAS \\ 1+(1-x)(1-x+x^2), & \text{if } \; \mathcal{K} = MKOBS \\ 1+(1-x)(1-\frac{x}{2}), & \text{if } \; \mathcal{K} = IFKOBS \end{cases} \tag{5}$$

**Lemma 1** *Let $Q, B$ be two $n \times n$ matrices such that $Q$ is orthogonal (i.e. $Q^TQ = QQ^T = I$). Then, for any expression of $f$ in* (5) *we have:*
*(i)* $f(QBQ^T) \;=\; Qf(B)Q^T$*;*
*(ii)* $f\left(diag(\gamma_1, \ldots, \gamma_n)\right) \;=\; diag(f(\gamma_1), \ldots, f(\gamma_n))$*.*

*Proof.* (i) If $f$ is a second degree polynomial, $f(x) = \alpha x^2 + \beta x + \gamma$ the proof is obvious. Let us then suppose that $f = \frac{2}{1+x}$. We obtain the equality in (i) by the following sequence of equalities (and also using that $Q^{-1} = Q^T$): $f(QBQ^T) = 2(I + QBQ^T)^{-1} = 2\left[Q(I+B)Q^T\right]^{-1} = Q\left[2(I+B)^{-1}\right]Q^T = Qf(B)Q^T$. For (ii) the proof is also obvious. ♢

Let now $Q^TA_0Q = D_0 = \text{diag}(\lambda_1^{(0)}, \ldots, \lambda_r^{(0)}, 0, \ldots, 0)$ be a spectral decomposition of $A$ (with $Q$ orthogonal), where $r = rank(A)$ and the eigenvalues $\lambda_i^{(0)} \in [0, 1), i = 1, \ldots, r$ (according to theorem 1). Then, for the algorithm $\mathcal{K}$-rhs a recursive argument gives us $A_k = QD_kQ^T$ where

$$D_k = \text{diag}(\lambda_1^{(k)}, \ldots, \lambda_r^{(k)}, 0, \ldots, 0), \;\; \lambda_i^{(k)} = f(\lambda_i^{(k-1)})\lambda_i^{(k-1)}, k \geq 1. \tag{6}$$

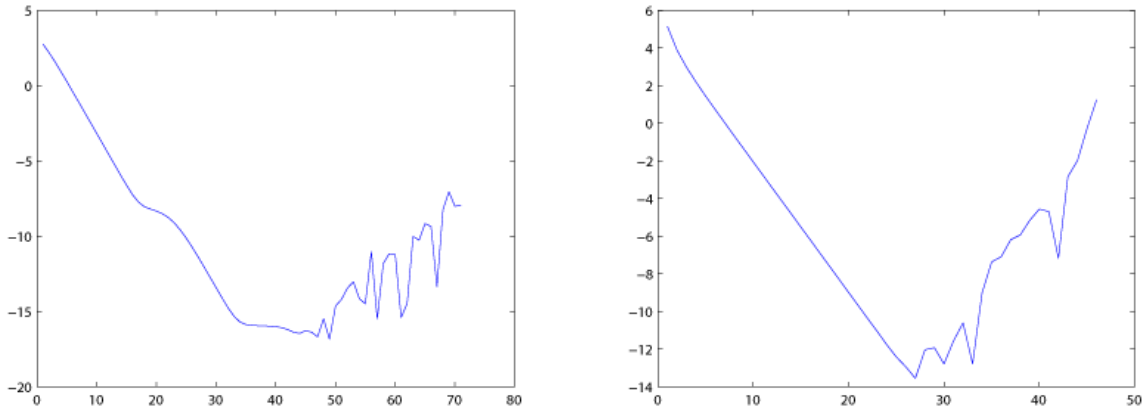

Figure 3: **P-pert; n = 32**: KOAS-rhs (left) and KOBS-rhs (right)

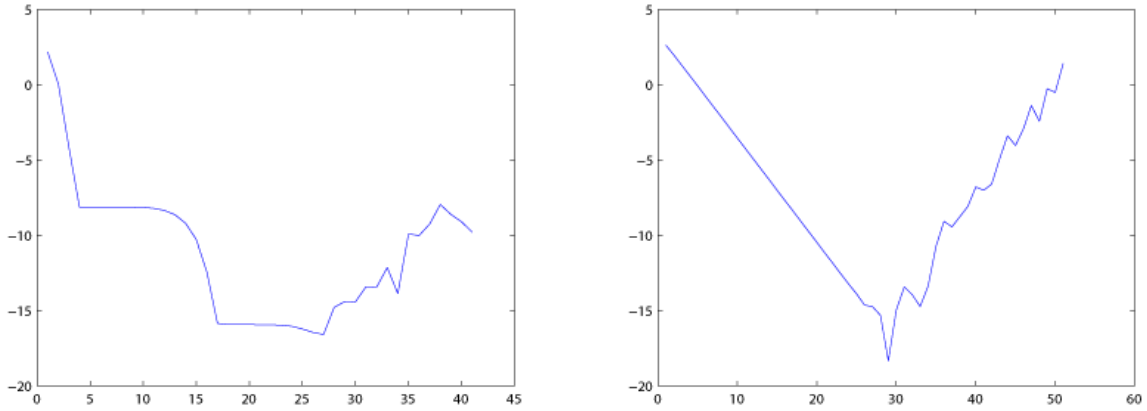

Figure 4: **P-pert; n = 32**: MKOBS-rhs (left) and IFKOBS-rhs (right)

The convergence analysis from theorem 1 is based on the analysis of the sequence implicitely generated by (6),

$$x_{k+1} = f(x_k)x_k, \quad x_0 \in \{\lambda_1^{(0)}, \ldots, \lambda_r^{(0)}\} \subset (0,1). \tag{7}$$

Additionally, it is shown that $x_k$ is strictly increasing to 1. Then, from (6) we get (from a "theoretical" point of view)

$$\lim_{k\to\infty} A_k = Q \text{ diag}\left(\lim_{k\to\infty} \lambda_1^{(k)}, \ldots, \lim_{k\to\infty} \lambda_r^{(k)}, 0, \ldots, 0\right) Q^T =$$

$$Q \text{ diag}\,(1, \ldots, 1, 0, \ldots, 0)\, Q^T = A^+A = P_{R(A)} = I - P_{N(A)}, \tag{8}$$

where $A^+$ is the Moore-Penrose pseudoinverse of $A$ and $P_{R(A)}, P_{N(A)}$ are the orthogonal projections onto the corresponding subspaces of $A$. Unfortunately, from a "practical" view point, the 0's on positions $r+1, \ldots, n$ in the initial diagonal matrix $D_0 = \text{diag}(\lambda_1^{(0)}, \ldots, \lambda_r^{(0)}, 0, \ldots, 0)$ are only $fl(0) =$"machine-zero" numbers (e.g. in double precision $fl(0) \approx 10^{-17}$). Then, for $\hat{x}_0 = fl(0)$, the iterative process (see (7)) $\hat{x}_{k+1} = f(\hat{x}_k)\hat{x}_k, k \geq 0$ generates larger and larger values. Thus, in practice, we start in algorithm $\mathcal{K}$-rhs, instead of $A$, with the matrix $\hat{A}_0 = Q\hat{D}_0Q^T$, with $\hat{D}_0 = \text{diag}(\lambda_1^{(0)}, \ldots, \lambda_r^{(0)}, fl(0), \ldots, fl(0))$. Then, after $k \geq 1$ iterations we get instead of $D_k$ from (6) the matrix

$$\hat{D}_k = \text{diag}(\lambda_1^{(k)}, \ldots, \lambda_r^{(k)}, \alpha_{r+1}(k), \ldots, \alpha_n(k)), \tag{9}$$

where the values $\alpha_j(k) > 0,\ j = r+1, \ldots, n$ can be large enough to destroy the computed approximations.

# 3 The regularization procedure

In order to overcome the difficulties mentioned in the above section, we shall formulate and try to solve the two problems from below.

**Problem P1.** *Find a treshold $\alpha^*$ such that, if*

$$\max\{\alpha_j(k),\ j=r+1,\ldots,n\}\geq\alpha^* \tag{10}$$

*then the value of $\hat{D}_k$ from (9) starts to affect the computations.*

**Problem P2.** *Suppose we know the above value $\alpha^*$; then find an integer $k^+(\alpha^*)\geq 1$ (as small as possible !) such that*

$$x_{k^+(\alpha^*)}\geq\alpha^*, \tag{11}$$

*where $x_{k^+(\alpha^*)}$ is the corresponding term of the sequence from (7) generated with the initial approximation*

$$x_0=10^{-17}. \tag{12}$$

We shall first try to solve the problem **P2**.

**Theorem 2** *A number with the property (11) – (12) is given by*

$$k^+(\alpha^*)=1+\left[\frac{\ln\left(\frac{(\alpha^*-x_0)H(\alpha^*)}{y_1}+1\right)}{\ln(1+\mathrm{H}(\alpha^*))}\right],\ \text{with } y_1=(f(x_0)-1)x_0 \tag{13}$$

*and the values of $H(\alpha^*)>0$ are as follows*

$$H(\alpha^*)\ =\ \begin{cases}\frac{1-2\alpha^*-{\alpha^*}^2}{(1+\alpha^*)^2}, & \text{if } \mathcal{K}=KOBS\\ 0.5-\alpha^*, & \text{if } \mathcal{K}=KOAS\\ 1-4\alpha^*, & \text{if } \mathcal{K}=MKOBS\\ 1-3\alpha^*, & \text{if } \mathcal{K}=IFKOBS\end{cases},\ \text{for } 0<\alpha^*<\frac{1}{4}. \tag{14}$$

*Proof.* For the sequence $(x_k)_{k\geq 0}$ defined by (7) with the initial approximation (12), let $k\geq 1$ be arbitrary, but fixed. We define $y_j=x_j-x_{j-1}, j=1,\ldots,k$ and obtain

$$x_k=x_0+y_1+\ldots+y_k. \tag{15}$$

On the other hand, from (7) we get $y_j=x_j-x_{j-1}=(f(x_{j-1})-1)\,x_{j-1}$, thus

$$y_j-y_{j-1}=$$

$$(f(x_{j-1})-1)\,x_{j-1}-(f(x_{j-2})-1)\,x_{j-2}=y_{j-1}g(x_{j-2},x_{j-1}), j=2,\ldots,k, \tag{16}$$

with $g(x_{j-2},x_{j-1})$ given by

$$g(x_{j-2},x_{j-1})=\begin{cases}\frac{1-x_{j-1}-x_{j-2}-x_{j-1}x_{j-2}}{(1+x_{j-1})(1+x_{j-2})}, & \text{if } \mathcal{K}=KOBS\\ \frac{1}{2}(1-x_{j-1}-x_{j-2}), & \text{if } \mathcal{K}=KOAS\\ 1+x_{j-1}^2+x_{j-2}^2-(x_{j-1}+x_{j-2}) & \\ (x_{j-1}^2+x_{j-2}^2+2-x_{j-1}-x_{j-2}), & \text{if } \mathcal{K}=MKOBS\\ 1-\frac{3}{2}(x_{j-1}+x_{j-2})+ & \\ \frac{1}{2}(x_{j-1}^2+x_{j-2}^2+x_{j-1}x_{j-2}), & \text{if } \mathcal{K}=IFKOBS\end{cases} \tag{17}$$

According to (11) we shall suppose that $0<x_0<x_1<\ldots<x_k<\alpha^*<\frac{1}{4}$. Then it can be easily proved that

$$g(x_{j-2},x_{j-1})>H(\alpha^*)>0,\forall j=2,\ldots,k, \tag{18}$$

with $H(\alpha^*)$ from (14), respectively. From (15) - (18) we then obtain

$$x_k>x_0+y-1+(1+H(\alpha^*))y_1+\ldots+(1+H(\alpha^*))^{k-1}y_1=$$

$$x_0+y_1\frac{[1+H(\alpha^*)]^k-1}{H(\alpha^*)}. \tag{19}$$

Table 1: Values of $k^+(\alpha^*)$

| Problem ( $\alpha^*$) | KOAS | KOBS | MKOBS | IFKOBS |
|---|---|---|---|---|
| P-cons ($10^{-3}$) | 80 | 48 | 47 | 47 |
| P-pert ($10^{-10}$) | 40 | 25 | 24 | 24 |

We then define $k^+(\alpha^*)$ as the smallest integer such that $\alpha^* < x_0 + y_1 \frac{[1+H(\alpha^*)]^k - 1}{H(\alpha^*)}$ and we successively get the values from (13) - (14) and the proof is complete. $\diamondsuit$

Now, we come back to the problem **P1**. Unfortunately, in this case we have no more a clear and complete solution as for problem **P2** (see the next section).

# 4 Numerical experiments

The solution to the problem **P1** essentially depends on the class of problems we solve. For example, in the case of discretizations of Fredholm first kind integral equations, we considered a smaller dimension (e.g. $n = 32$ for our problems **P-cons** and **P-pert**) and made several tests which gave us at the end an appropriate choice for the treshold $\alpha^*$, namely: $\alpha^* = 10^{-3}$ for **P-cons** and $\alpha^* = 10^{-10}$ for **P-pert**. For these values we computed the corresponding integers $k^+(\alpha^*)$ defined by (13) and we got the values from Table 1.

Then, we performed experiments on both problems, with $n = 32, 64, 128, 256, 512$ and we obtained the results in figures 5 - 8. We can see there that the values of $k^+(\alpha^*)$ determined in the particular case $n = 32$ are good also for bigger dimensions, in the sense that before it no instability appear in the computed solution.

REMARK 1. *The big difference between $\alpha = 10^{-3}$ for* **P-cons** *and $\alpha = 10^{-10}$ for* **P-pert** *can be explained by using the information given in theorem 1(ii). According to this, in the inconsistent case we have $\lim_{k\to\infty} \| b^k \| = +\infty$. Moreover, e.g. in the case of the algorithm IFKOBS algorithm we get $b^k = A_k x + 2^k P_{N(A)}(b)$, where $x \in \mathbb{R}^n$ is a vector with the property $Ax = P_{R(A)}(b)$; then, $\lim_{k\to\infty}(A_k x) = (\lim_{k\to\infty} A_k)x = (I - P_{N(A)})x$. Thus, $\| b^k \|$ "goes" to $+\infty$ as $2^k \| P_{N(A)}(b) \|$ does and, if we let the algorithm to "run", e.g. until $x_k$ will exceed the value $\alpha = 10^{-3}$, the value $2^k \| P_{N(A)}(b) \|$ will become too large and will pollute the vector $b^k$ and so, the approximation $A_k b^k$ of $x_{LS}$.*

# References


[1] H. W. Engl, *Regularization methods for the stable solution of inverse problems*, Surv. Math. Ind., 3(1993), pp. 71–143.

[2] Z. Kovarik, *Some iterative methods for improving orthogonality*, SIAM J. Numer. Anal., 7(3)(1970), pp. 386–389.

[3] M. Mohr, C. Popa, and U. Rüde, *An iterative algorithm for approximate orthogonalization of symmetric matrices*, Intern. J. Computer Math., 81(2)(2004), pp. 215–226.

[4] M. Mohr, and C. Popa, *Numerical solution of symmetric least-squares problems with an inversion-free Kovarik type-algorithm*, Intern. J. Computer Math., 85(2)(2008), pp. 271–286.

[5] C. Popa, *Extension of an approximate orthogonalization algorithm to arbitrary rectangular matrices*, Linear Alg. Appl., 331(2001), pp. 181–192.

[6] C. Popa, *A method for improving orthogonality of rows and columns of matrices*, Intern. J. Computer Math., 77(3)(2001), pp. 469–480.

[7] C. Popa, *On a modified Kovarik algorithm for symmetric matrices*, Annals of Ovidius University Constanta, Series Mathematics, XI(1)(2003), pp. 147–156.

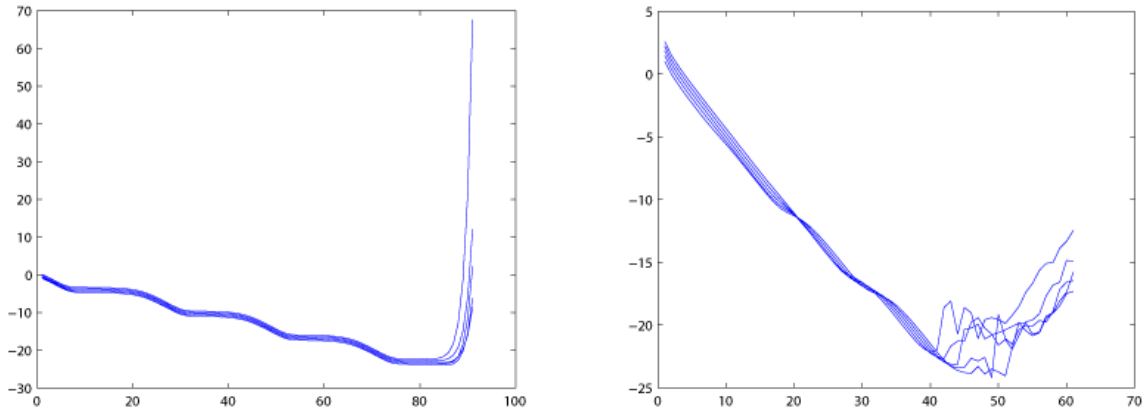


Figure 5: **P-cons; n = 32, 64, 128, 256, 512**: KOAS-rhs (left) and KOBS-rhs (right)

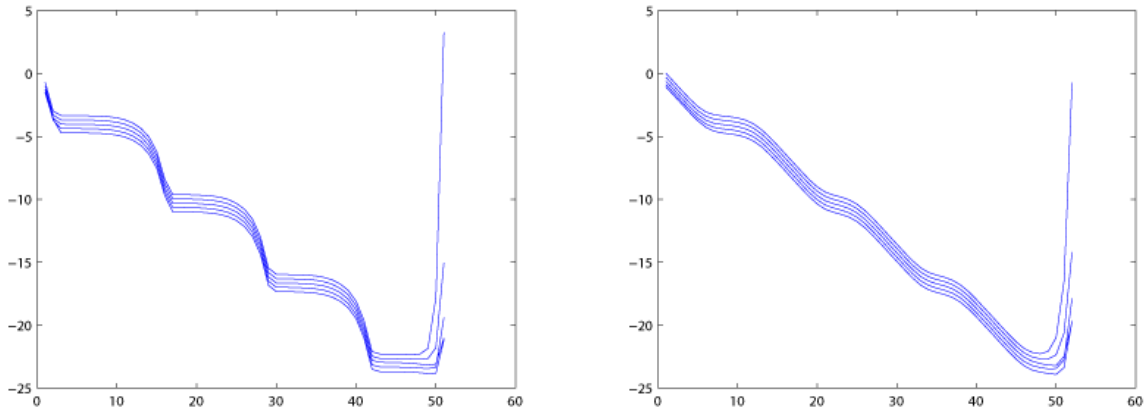


Figure 6: **P-cons; n = 32, 64, 128, 256, 512**: MKOBS-rhs (left) and IFKOBS-rhs (right)

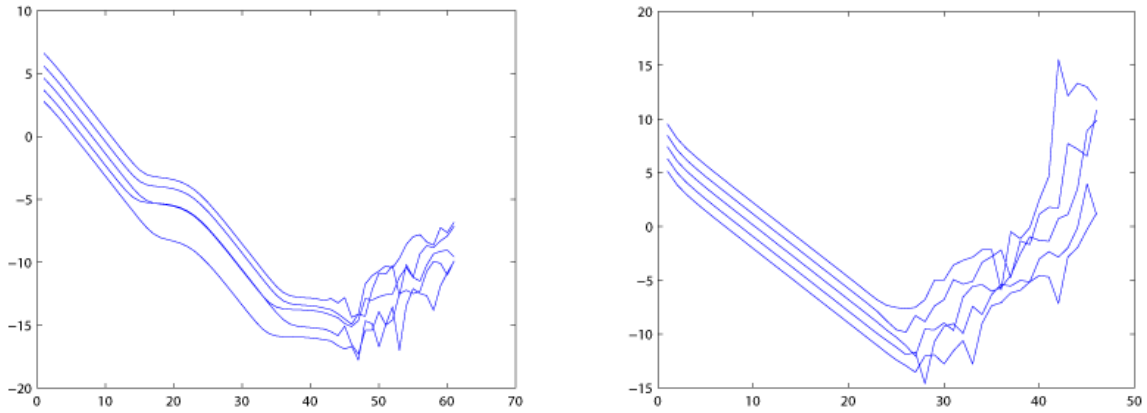


Figure 7: **P-pert; n = 32, 64, 128, 256, 512**: KOAS-rhs (left) and KOBS-rhs (right)

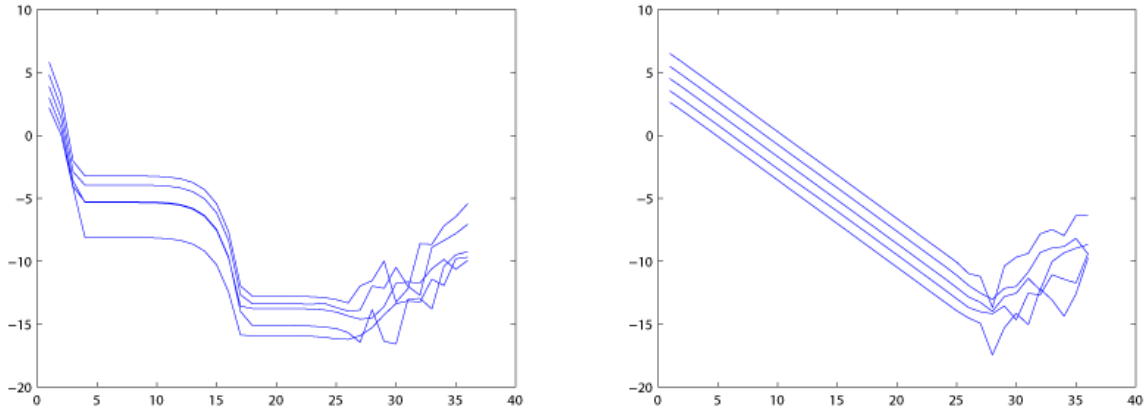


Figure 8: **P-pert; n = 32, 64, 128, 256, 512**: MKOBS-rhs (left) and IFKOBS-rhs (right)