

APPLICATION OF KERNEL RIDGE REGRESSION TO NETWORK LEVELLING VIA *MATHEMATICA*

L. Völgyesi (1), B. Paláncz (2), K. Fekete (3), Gy. Popper (4)

Budapest University of Technology and Economics, H-1521 Budapest, Műegyetem rkp
3

(1) Department of Geodesy and Surveying, Research Group for HAS-BUTE Physical Geodesy and Geodynamics, E-mail: volgyesi@eik.bme.hu

(2) Department of Photogrammetry and Geoinformatics, E-mail: palancz@epito.bme.hu

(3) Department of Photogrammetry and Geoinformatics, E-mail: kfekete@epito.bme.hu

(4) Department of Structural Mechanics, Research Group for HAS-BUTE Computational Structural Mechanics, E-mail: gypopper@epito.bme.hu

ABSTRACT

A new method based on support vector regression (SVR) has been developed for network levelling. Employing zero insensitive margin and first order polynomial kernel, the general form of SVR has been reduced to a kernel ridge regressor, which is a linear function approximator. Then this function approximation problem can be transformed into an adjustment problem, simply using proper recasting of the variables. Only one part of the measured values (training equations) is considered in the adjustment, the other part of them (test equations) is used to compute the risk of the data generalization. Then the quality of the estimation can be measured by computing the performance index of the levelling, a value which is a trade off between adjustment quality (residual of the test equations) and the adjustment risk (the ratio of the residual of the test equations and that of the training equations). This performance index can be optimized with the regularization term of the ridge regressor. The algorithm was implemented in *Mathematica* 5.1 and demonstrated by numerical example.

Keywords: adjustment, levelling, support vector regression, kernel ridge regression, software *Mathematica* 5.1

INTRODUCTION

Robust adjustment methods are well developed in geoscience application, (Jianjun, 1996). These methods are based on different statistical models, and their common feature is that they are robust against the influence of gross errors, however their can be distinguished according to the type of robustness they being for, like distribution robustness, data robustness, model robustness and so on, (Wicki, 1999).

The reliability of adjustment was first introduced by Baarda (Wicki, 1991, Carosio, 1996). He considered three factors, which may have influence on the reliability of the adjustment process: the structure of the geodetic network, the error modelling and the statistical test of measured data and computed values.

These approaches use all of the measured data for adjustment analysis, following the famous saying, that "one measurement is not a measurement". However this saying may suggest a different approach, especially on the light of the artificial learning theory. Artificial learning methodology uses training data in order to select a hypothesis from the hypothesis space. The quality of learning, the consistency of the selected hypothesis, can not be tested on the training set, while hypothesis that is consistent on this set, may not correct on unseen data. Therefore, the so called data generalization ability is a key factor, (Mitchell, 1997).

In this paper, we apply this principle to geodetic network adjustment, using only one part of the measurements for the adjustment, while the other part of them will serve for testing. Kernel ridge regression is employed in order to make a trade-off between the quality and the risk of the adjustment and in this way the overall performance index of the adjustment can be optimized.

1. ADJUSTMENT PROBLEM AND ITS TRADITIONAL SOLUTION

First, we shall demonstrate the method with a very simple levelling problem, see Fig. 1.

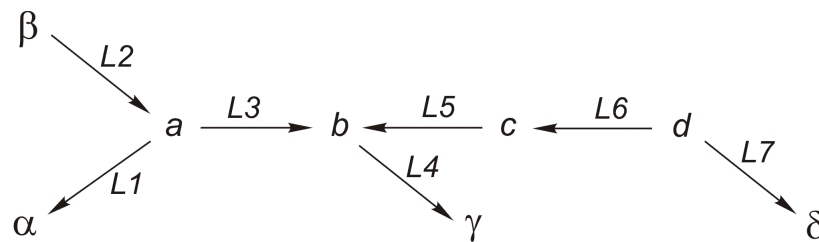


Fig. 1. A simple levelling problem

The unknown values of heights to be estimated via adjustment are, a , b , c and d ; values of α , β , γ and δ are the known heights; and L_i are the measured height differences between points. Arrows show the direction of ascents. The adjusting equations can be written according to Fig.1

$$\text{eqs}=\{L1==\alpha-a, L2==a-\beta, L3==b-a, L4==\gamma-b, L5==b-c, L6==c-d, L7==\delta-d\};$$

Now, we transform these equations into matrix form. The matrix of the system is

$$\mathbf{A}=\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where the system matrix has m rows and n columns

$$\{m, n\}=\text{Dimensions}[\mathbf{A}]$$

$$\{7, 4\}$$

and $m > n$.

The measured values are

data={L1→0.351, L2→1.942, L3→0.863, L4→0.631, L5→1.798, L6→0.673,
L7→1.107, L8→1.124, α→720.301, β→718.006, γ→721.494, δ→719.567,
ε→719.566};

Then the right handside vector is

$\mathbf{b}=\{719.95, -719.948, -0.863, 720.863, -1.798, -0.673, 718.46\}$

with dimension m .

The rank of A is

MatrixRank[A]

4

and the rank of the augmented matrix is

$$\mathbf{Ab} = \begin{pmatrix} 1 & 0 & 0 & 0 & 719.95 \\ -1 & 0 & 0 & 0 & -719.948 \\ 1 & -1 & 0 & 0 & -0.863 \\ 0 & 1 & 0 & 0 & 720.863 \\ 0 & -1 & 1 & 0 & -1.798 \\ 0 & 0 & -1 & 1 & -0.673 \\ 0 & 0 & 0 & 1 & 718.46 \end{pmatrix}$$

MatrixRank[Ab]

5

consequently the system is overdetermined, inconsistent having no solution. However, solution can be seeked in the least square sense, employing pseudoinverse of A

PI=PseudoInverse[A];

$$\begin{pmatrix} \frac{7}{18} & -\frac{7}{18} & \frac{2}{9} & \frac{1}{6} & -\frac{1}{18} & -\frac{1}{18} & \frac{1}{18} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{9} & -\frac{1}{9} & -\frac{2}{9} & \frac{1}{3} & \frac{5}{9} & -\frac{4}{9} & \frac{4}{9} \\ \frac{1}{18} & -\frac{1}{18} & -\frac{1}{9} & \frac{1}{6} & \frac{5}{18} & \frac{5}{18} & \frac{13}{18} \end{pmatrix}$$

then the solution is

$\mathbf{x}=\mathbf{PI} \cdot \mathbf{b}$

{719.964, 720.857, 719.084, 718.435}

2. QUALITY AND RISK OF THE ADJUSTMENT

The *quality* of the least square adjustment can be characterized by the average residual, namely the total residual divided by the number of rows of A ,

$q = (1/m) \text{Norm}[A \cdot \mathbf{x} - \mathbf{b}]$

0.00809776.

Individual residual differing considerably from this average may indicate outlier measurement then the correspondig row should be eliminated from the system. However in geoinformatics it frequently happens that data collections are irregular and far from substantial for adjustment. This fact could lead to overadjustment, a poor data gener-

alization, which means that we get small average and individual residiums for the rows of matrix A , but large average residiums for adjustment equations are not considered as rows in A .

To characterize data generalization of the adjustment method, the adjustment equations can be divided into two groups, equations for training (adjusting) and equations for testing.

Let us consider the first five equations (rows) in our original system, namely we carry out the adjustment of $A^+x^+=b^+$

$$A^+ = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$b^+ = \{-0.863, 720.863, -1.798, -0.673, 718.46\}.$$

The rank of the matrix of the training equations should be also m

`MatrixRank[A+]`

4

$$Ab^+ = \begin{pmatrix} 1 & -1 & 0 & 0 & -0.863 \\ 0 & 1 & 0 & 0 & 720.863 \\ 0 & -1 & 1 & 0 & -1.798 \\ 0 & 0 & -1 & 1 & -0.673 \\ 0 & 0 & 0 & 1 & 718.46 \end{pmatrix}$$

`MatrixRank[Ab+]`

5

The corresponding solution is

$$x^+ = \text{PseudoInverse}[A^+] \cdot b^+ \\ \{720.017, 720.88, 719.099, 718.443\}$$

the average residual, q^+ is

`m+=Length[b+]`

5

`q+=(1/m+) Norm[A+ . x+ - b+]`

0.0068

Let us test our solution with the remained two equations (test equations), which were left out from the adjustment process,

$$A^- = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

the corresponding right side

$$b^- = \{719.95, -719.948\}$$

The average residium of the test equations, q^- is

`m-=Length[b-]`

2

`q-=(1/m-) Norm[A- . x+ - b-]`

0.0480885

We can characterize the risk of the generalization of our solution with the following heuristic function of the ratio of the two average residiums, $\xi = q^- / q^+$

$$r[\xi] = \text{Max}[\xi - 1, 0];$$

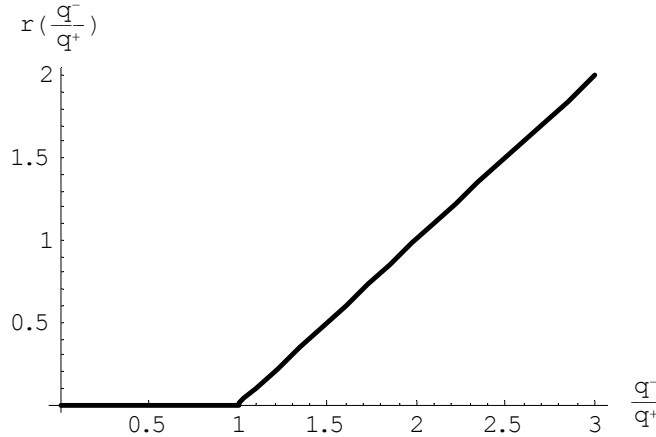


Fig. 2. Risk of adjustment

This means that our solution has a risk from the point of view of data generalization, when the residium of the test equations is higher than the residium of the adjustment itself. Higher this ratio, higher the risk is. With other words, in case of high risk, there is a chance, that if there would be more equations from further measurements at our disposal, the result of the adjustment would change considerably.

In our case this risk is now

$$r[q^- / q^+] // N$$

6.07183

which can be considered a high risk.

Using traditional adjustment technique (pseudoinverse), there is no free parameter to influence these residiums consequently we have no means to reduce the risk of data generalization.

In this paper, we suggested a technique based on kernel ridge regression, which tackles this problem. The first step is to transform the adjustment problem into a function approximation problem.

3. ADJUSTMENT AS KERNEL RIDGE REGRESSION

Adjustment problems can be written in the following form

$$\begin{aligned} & \text{minimise } \sum_{i=1}^m \xi_i^2 \\ & \text{subject to } \xi_i = \sum_{j=1}^n A_{ij} x_j - b_i \quad i = 1, \dots, m \end{aligned}$$

(x is the vector of unknowns, A is the coefficient matrix of the observation equations and b is the vector of absolute terms) which can be transformed into a kernel ridge regression problem employing the primal problem of support vector regression with

zero insensitive margin, $\epsilon = 0$, also ignoring bias term, (Cristianini et al, 2003). This yields

$$\begin{aligned} & \text{minimise } \lambda(\|w\|)^2 + \sum_{i=1}^m \eta_i^2 \\ & \text{subject to } \eta_i = y_i - \sum_{j=1}^n w_j \alpha_{ij} = y_i - \langle w, \alpha_i \rangle, \quad i = 1, \dots, m \end{aligned}$$

where λ is the regularization parameter. Using the following recasting of the variables,

$$y_i = b_i, \quad \alpha_{ij} = A_{ij}, \quad w_j = x_j, \quad j = 1, \dots, n, \quad i = 1, \dots, m$$

for which, we derive the Lagrangian,

$$\text{minimise } L(w, \eta, \rho) = \lambda(\|w\|)^2 + \sum_{i=1}^m \eta_i^2 + \sum_{i=1}^m \rho_i (y_i - \langle w, \alpha_i \rangle - \eta_i)$$

with respect w , where ρ is the Lagrangian multiplier. Differentiating, and imposing stationarity, we obtain that

$$w = \frac{1}{2\lambda} \sum_{i=1}^m \rho_i \alpha_i, \quad \text{and} \quad \eta_i = \frac{\rho_i}{2}.$$

Resubstituting these relations gives the following dual problem,

$$\text{maximise } W(\rho) = \sum_{i=1}^m y_i \rho_i - \frac{1}{4\lambda} \sum_{i,j=1}^m \rho_i \rho_j \langle \alpha_i, \alpha_j \rangle - \frac{1}{4} \sum_{i=1}^m \rho_i^2$$

that for convenience we rewrite in vector form,

$$W(\rho) = y^T \rho - \frac{1}{4\lambda} \rho^T M \rho - \frac{1}{4} \rho^T \rho$$

differentiating with respect ρ and imposing stationarity,

$$-\frac{1}{2\lambda} M \rho - \frac{1}{2} \rho + y = 0$$

which is giving the solution

$$\rho = 2\lambda(M + \lambda I)^{-1} y$$

where M denotes the Gram matrix $M_{ij} = \langle \alpha_i, \alpha_j \rangle$.

Now, we can express the optimal linear function as

$$f(x) = \langle w, \kappa \rangle = \left\langle \frac{1}{2\lambda} \sum_{i=1}^m \rho_i \alpha_i, \kappa \right\rangle = y^T (M + \lambda I)^{-1} k$$

where κ is the formal independent variable in the original KRR problem and

$$k_i = \langle \alpha_i, \kappa \rangle.$$

Using the terms of the adjustment, see casting, we get

$$f(x) = \langle x, \kappa \rangle = x_1 \kappa_1 + x_2 \kappa_2 + \dots + x_n \kappa_n = b^T (M + \lambda I)^{-1} k$$

where I stands for the unit matrix; and the Gram matrix, M is,

$$M_{ij} = \langle A_i, A_j \rangle, \quad i, j = 1, \dots, n$$

and

$$k_i = \langle A_i, \kappa \rangle$$

where A_i is the i -th row in matrix A .

4. SOLUTION VIA KERNEL RIDGE REGRESSION

For linear adjustment problem we employ polynomial kernel of first order, (Schölkopf et al, 1999),

$$K[u, v] := u \cdot v$$

The Gram matrix for the training equations is

$$M = \text{Table}[K[A^+[[i]], A^+[[j]]], \{i, 1, m^+\}, \{j, 1, m^+\}];$$

$$\begin{pmatrix} 2 & -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 \\ 1 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

and

$$k = \{\kappa_1 - \kappa_2, \kappa_2, -\kappa_2 + \kappa_3, -\kappa_3 + \kappa_4, \kappa_4\}$$

Then the regression function, with regularization parameter $\lambda = 10^{-5}$

$$\lambda = 0.00001;$$

$$f = b^+ \cdot (\text{Inverse}[M + \lambda \text{IdentityMatrix}[m^+]] \cdot k)$$

$$719.994 \quad \kappa_1 + 720.864 \quad \kappa_2 + 719.081 \quad \kappa_3 + 718.43 \quad \kappa_4$$

The solution of the adjustment problem is,

$$x^+ = \{719.994, 720.864, 719.081, 718.43\}$$

Computing the average residual of the adjustment, we get

$$qR^+ = (1/m^+) \text{Norm}[A^+ \cdot x^+ - b^+]$$

$$0.00815105$$

This is somewhat higher value than the value resulted from pseudoinverse solution, q^+

$$q^+$$

$$0.0068$$

Let us test our solution with the remained two equations. The average residium of the test equations is

$$qR^- = (1/m^-) \text{Norm}[A^- \cdot x^- - b^-]$$

$$0.0315464$$

the risk of the adjustment is

$$r[qR^-/qR^+]/N$$

$$2.87023$$

which is lower value than that of the pseudoinverse solution,

$$r[q^-/q^+]$$

$$6.07183$$

The quality of the adjustment, $q^+ + q^-$ in case of pseudoinverse solution,

$$(q^+ + q^-)$$

0.0548885

and employing kernel ridge regression, we get also better result namely

(qR^+/qR^-)
0.0396975.

This means, somewhat less "tight fitting" adjustment of the training set, provides smaller residium of the test set. Using this type of technique for adjustment, one can make a trade-off between adjustment quality and adjustment risk. This balance can be controlled by the regularization parameter of the kernel ridge regression. Increasing this λ parameter, the risk can be reduced, but the adjustment quality will be worse.

5. FINDING OPTIMAL REGULARIZATION PARAMETER

Let us consider the following objective function

$$Q(\lambda, \mu) = (qR^+ + qR^-) + \mu r \left(\frac{qR^-}{qR^+} \right)$$

The influence of the risk can be controlled by this weighting parameter μ . We may consider this function as a performance index of the adjustment, namely smaller is the value of $Q(\lambda, \mu)$, more efficient the adjustment is. This performance index for $\lambda = 0.00001$ and $\mu=1$, is

$Q = (qR^+ + qR^-) + r [qR^- / qR^+]$
2.90992.

While the value of this performance index for the traditional pseudoinverse solution is,

$(q^+ + q^-) + r [q^- / q^+]$
6.12672.

The optimal regularization parameter, λ can be determined as the solution of the following optimization problem,

$$\underset{\lambda}{\text{minimize}} Q(\lambda, \mu).$$

We just mention here, that it would be also possible to generalize this objective function using different weighting factors for quality and risk.

6. IMPLEMENTATION IN *MATHEMATICA*

Let us summarize the steps of the computation as a *Mathematica* function.

```
KernelRidgeRegression[ $\lambda$ _, A1_, b1_, A2_, b2_,  $\mu$ _]:=Module[{K, u, v,
M, i, j, m1, m2, n, k, f, x1, qR1, qR2, q, rR, r, QR,  $\xi$ },
{m1, n}=Dimensions[A1]; {m2, n}=Dimensions[A2];
K[u_, v_]:=u.v; r[ $\xi$ _]:=Max[ $\xi$ -1., 0.];
M=Table[K[A1[[i]], A1[[j]]], {i, 1, m1}, {j, 1, m1}];
k=Table[A1[[i]].Table[kj, {j, 1, n}], {i, 1, m1}];
f=Simplify[b1.(Inverse[M+ $\lambda$  IdentityMatrix[m1]].k)];
x1=Flatten[Table[Coefficient[f, kj, {i, 1, n}]];
qR1=(1/m1) Norm[A1.x1-b1]; qR2=(1/m2) Norm[A2.x1-b2];
q=qR1+qR2; rR=r[qR2/qR1]; QR=q+ $\mu$ .rR;
{QR, rR, q, qR1, qR2, x1}];
```


Let us check it,

```
KernelRidgeRegression[0.00001, A+, b+, A-, b-, 1]
{2.90992, 2.87023, 0.0396975, 0.00815105, 0.0315464,
 {719.994, 720.864, 719.081, 718.43}}
```

which is the same, what we computed "zu Fuß",

```
{Q, r[qR-/qR+], (qR++qR-), qR+, qR-, x+}
{2.90992, 2.87023, 0.0396975, 0.00815105, 0.0315464,
 {719.994, 720.864, 719.081, 718.43}}
```

Let us plot the performance index as function of the regularization parameter, λ

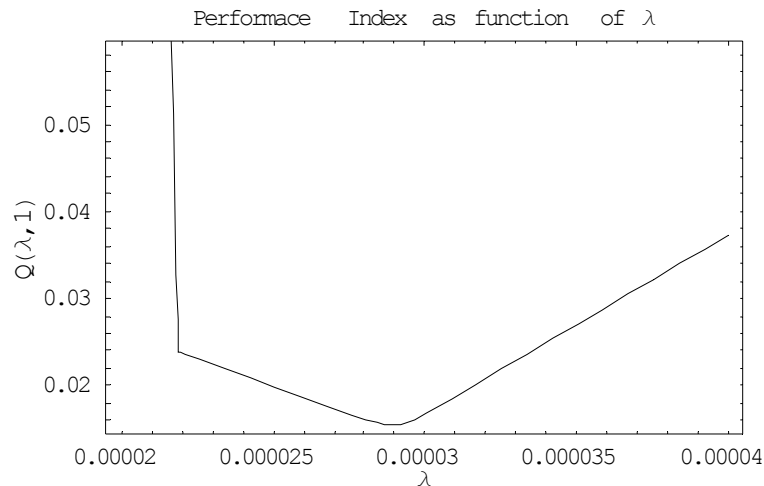


Fig. 3. The performance index, $Q(\lambda, 1)$ function

The optimal regularization is

```
λopt=0.0000289582
```

and for the optimal solution, we get

```
sol=KernelRidgeRegression[λopt, A+, b+, A-, b-, 1]
{0.0154122, 0., 0.0154122, 0.0146837, 0.000728537,
 {719.949, 720.833, 719.047, 718.407}}
```

the performance index,

```
sol[[1]]
0.0154122.
```

The Kernel Ridge Regression (KRR) method gives the following adjustment result

```
x+=sol[[6]]
{719.949, 720.833, 719.047, 718.407}
```

the residual of the all equations

```
(1/m)Norm[A.x+-b]
0.0104904
```

while employing the traditional pseudoinverse for all equations

```
x
{719.964, 720.857, 719.084, 718.435}
```

residual for traditional solution:

```
(1/m)Norm[A.x-b]
```

0.00809776 .

The average residual of the traditional solution is smaller than that of the KRR method, but this solution gives no information about adjustment risk, while the second one is able to indicate this risk and even provide an optimal trade off between risk and adjustment quality.

In case of using all equations for training, *we have no information about adjustment risk*, and only the quality of the adjustment can be computed, because *no test equations are available*. Employing KRR we loose somewhat in quality but gain information about the risk of data generalization, and find the optimal trade off between quality and risk. KRR method uses a part of information being at our disposal to estimate and increase the reliability of the adjusted values and at same time tries to preserve adjustment quality.

7. GEODETIC APPLICATION

Now let us employ the KRR method for levelling the following synthetic network.

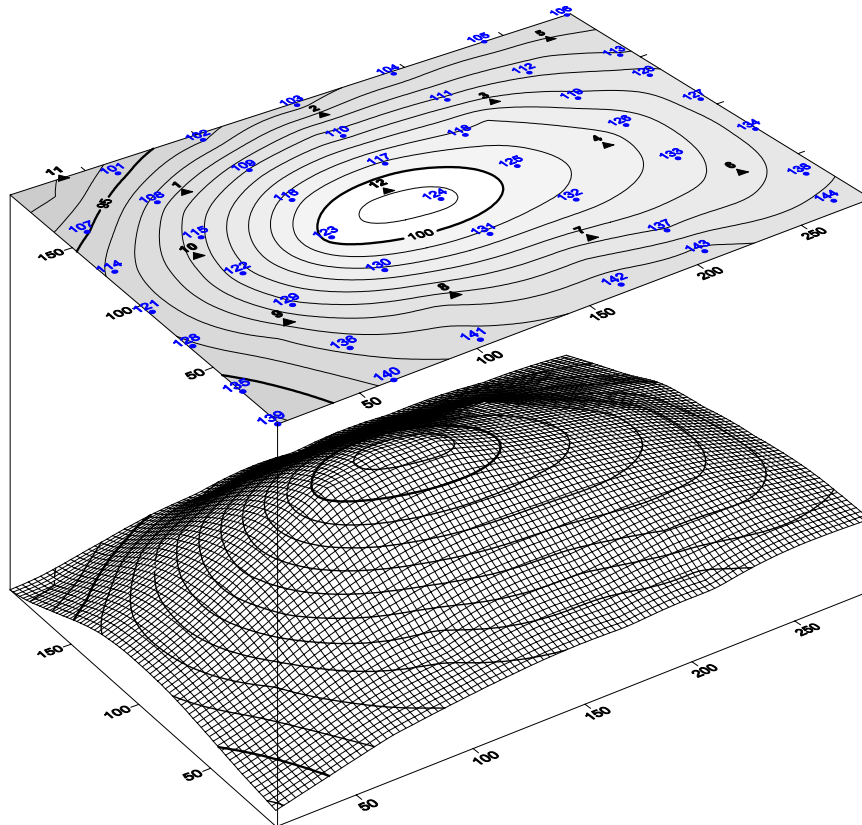


Fig. 4. The terrain model for the test computations

In Fig. 5 triangles indicate points where values of height are known. Arrows show the direction of ascents. The thicker lines indicate the testing relations (equations). The total number of equations is 115. We used 90 equations for training and 25 equations for testing. For example the 16th equation looks like

$$\begin{aligned} \text{Eqs1} & [[16]] \\ -h_5 + h_{106} & == -0.88 \end{aligned}$$

similarly the 3rd test equation is,

Eqs2[[3]]
 $h_{103}-h_{104}==0.141$.

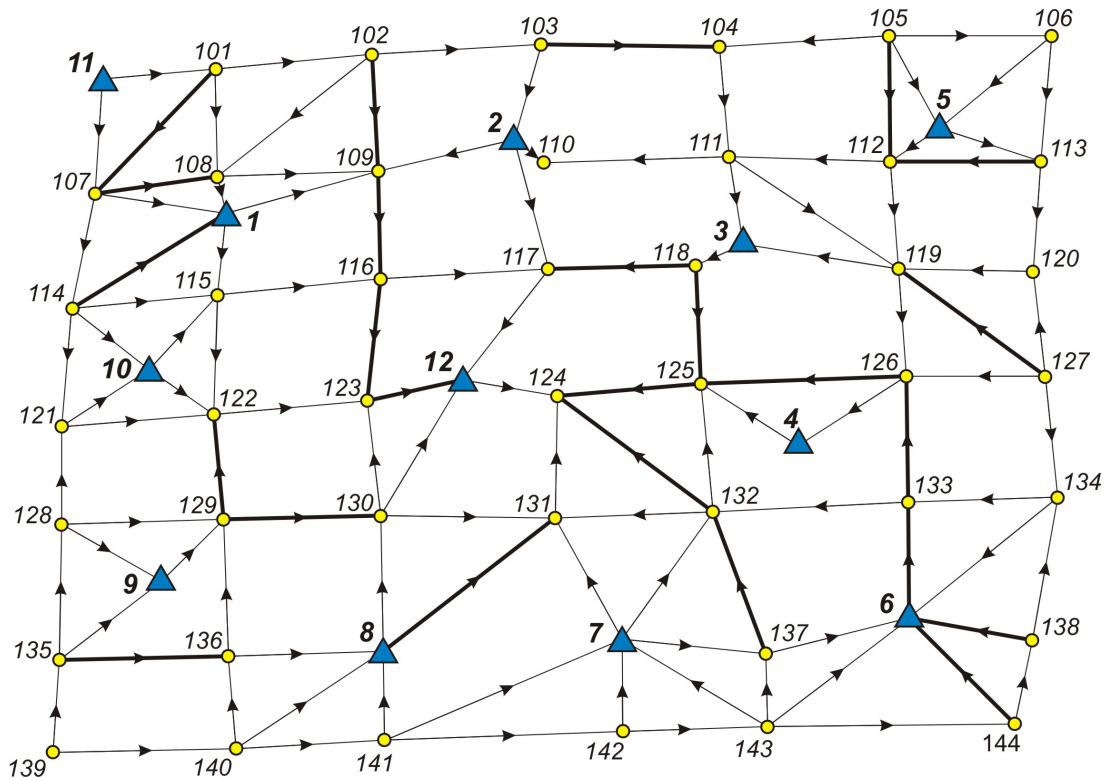


Fig. 5. The levelling network for the test computations

The value of height is known in twelve points. They are

$\{h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9, h_{10}, h_{11}, h_{12}\}$.

The values,

$\{96.844, 97.439, 99.002, 99.332, 96.996, 98.252, 97.637, 6.743, 96.005, 96.744, 93.981, 100.684\}$;

The unknown 44 variables of levelling, where the heights should be computed are

$H=\{h_{101}, h_{102}, h_{103}, h_{104}, h_{105}, h_{106}, h_{107}, \langle\langle 31 \rangle\rangle, h_{139}, h_{140}, h_{141}, h_{142}, h_{143}, h_{144}\}$.

Let us convert these equations into matrix form. For the training equations, we get

```
Axb1=LinearEquationsToMatrices[Eqs1,H];
A1=Axb1[[1]];
b1=Axb1[[2]];
{m1,n}=Dimensions[A1]
{90,44}.
```

For the test equations, we get

```
Axb2=LinearEquationsToMatrices[Eqs2,H];
A2=Axb2[[1]];
b2=Axb2[[2]];
{m2,n}=Dimensions[A2]
{25,44}.
```

Employing KRR method, the performance index as function of λ , in case low level risk influence, $\mu = 0.00005$

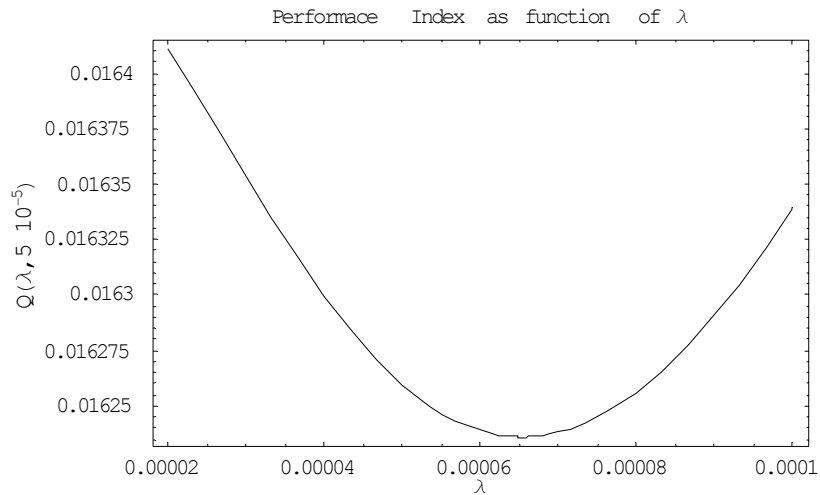


Fig. 5. Performance index as function of the regularization parameter, λ

and the optimal value,

$$\lambda_{opt}=0.00006537.$$

Then the KRR method gives the following values,

$$\text{sol}=\text{KernelRidgeRegression}[\lambda_{opt}, \mathbf{A1}, \mathbf{b1}, \mathbf{A2}, \mathbf{b2}, \mu];$$

the optimal performance index,

$$0.016236,$$

the levelled values are

$$\mathbf{x1}=\{94.812, 95.440, 95.787, 96.096, 96.093, 96.107, 94.967, \\ 96.243, 97.560, 97.831, 97.641, 97.353, 97.027, 95.678, 97.441, \\ 99.304, 99.613, 99.130, 98.143, 97.668, 95.697, 97.865, 100.245, \\ 100.693, 99.728, 98.822, 97.667, 95.437, 97.397, 98.881, 99.323, \\ 99.300, 98.702, 97.722, 94.775, 96.145, 97.948, 97.699, 94.405, \\ 95.671, 96.297, 96.606, 96.802, 97.302\}$$

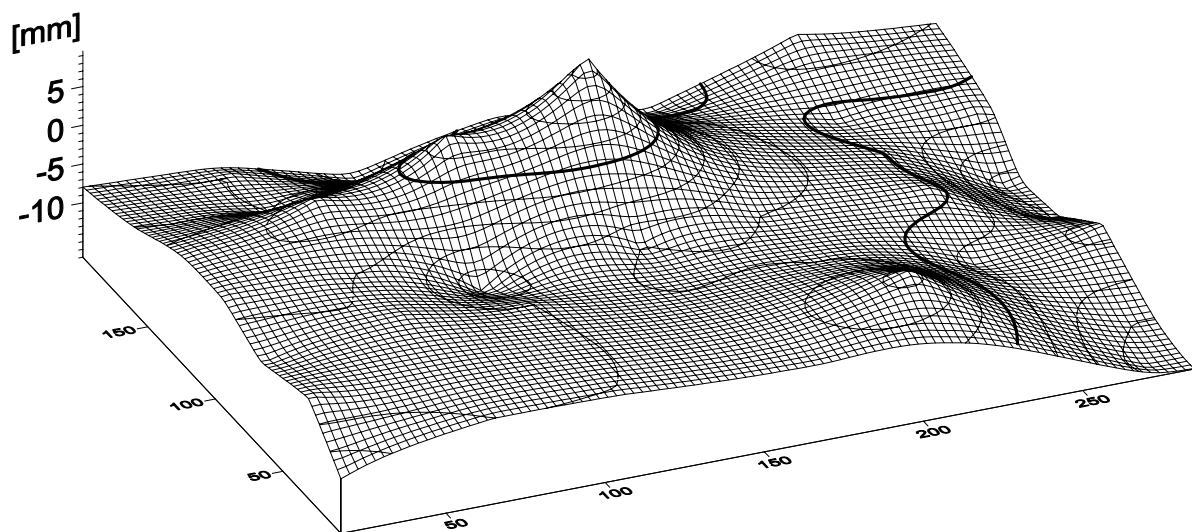


Fig. 6. Errors of the computed heights

Computed and original heights were compared and standard error was computed by the formula

$$m_x = \sqrt{\frac{\sum_{i=1}^{44} (x_{1_i} - x_i)^2}{44}}$$

where x_i are the original and x_{1_i} are the computed heights. The standard error is ± 9 mm. Error distribution of computed heights for the terrain model can be seen in Fig. 6.

8. CONCLUSIONS

Artificial learning methodology uses training data in order to select a hypothesis from the hypothesis space. The quality of learning, the consistency of the selected hypothesis, can not be tested on the training set, while hypothesis that is consistent on this set, may not correct on unseen data. Therefore, the so called data generalization ability is a key factor. We have applied this principle to levelling network adjustment, using only one part of the measurements for the adjustment, while the other part of them will serve for testing. Kernel ridge regression is employed in order to make a trade-off between the quality and the risk of the adjustment and in this way the overall performance index of the adjustment can be optimized. In geoinformatics it frequently happens that data collections are irregular and far from substantial for adjustment. This fact could lead to overadjustment, a poor data generalization, which means that we get small average and individual residiums for the rows of matrix A , but large average residiums for adjustment equations are not considered as rows in A . Somewhat less "tight fitting" adjustment of the training set, provides smaller residium of the test set. Using this type of technique for adjustment, one can make a trade-off between adjustment quality and adjustment risk. This balance can be controlled by the regularization parameter of the kernel ridge regression. Increasing this regularization parameter, the risk can be reduced, but the adjustment quality will be worse.

ACKNOWLEDGEMENTS

We should thank for the funding of the above investigations to the National Scientific Research Fund OTKA T-037929 and T-37880.

REFERENCES

- Carosio A. (1996): Fehlertheorie... ETH Zürich, Dep. Geodätische Wissenschaften.
Cristianini N, Shawe-Taylor J. (2003): An Introduction to Support Vector Machines and other kernel-based learning methods, Cambridge University Press.
Jianjun Z. (1996): Robustness and robust estimate, *Journal of Geodesy*, pp.586-590.
Mitchell T. (1997): Machine Learning, McGraw-Hill.
Schölkopf B, Smola A. J, Müller K. (1999): Kernel principal... In Schölkopf B, et all. Ed. Advances in Kernel Methods – Support Vector Learning, pp. 327-352, MIT Press.
Wicki F. (1991): Zuverlässigkeitstheorie, Beurteilungskriterien... ETH Zürich, Institute für Geodäsie und Photogrammetrie, Bericht Nr. 176.

Wicki F. (1999): Robuste Schätzverfahren... ETH Zürich, Institute für Geodäsie und Photogrammetrie, Dissertation ETH Nr. 12894.

* * *

Völgyesi L, Palánc B, Fekete K, Popper Gy. (2005) [Application of Kernel Ridge Regression to Network Levelling via Mathematica](#). Reports on Geodesy, Warsaw University of Technology, Vol. 73, Nr. 2. pp. 263-276.

Dr. Lajos VÖLGYESI, Department of Geodesy and Surveying, Budapest University of Technology and Economics, H-1521 Budapest, Hungary, Műegyetem rkp. 3.
Web: <http://sci.fgt.bme.hu/volgyesi> E-mail: volgyesi@eik.bme.hu