DESIGN OF SPECIAL PURPOSE NETWORKS BY RISK REDUCTION

T. Bouloucos
International Institute for Aerospace Survey and Earth Sciences (ITC)
350, Boulevard 1945, 7511 AL ENSCHEDE
The Netherlands
Commission III/1

1. Prologue

The design strategy which is developed in this paper is aimed at special purpose photogrammetric blocks, geodetic or combined networks. Special purpose blocks or nets are those in which the coordinates of the network points are determined in a coordinate system which has been specified for a particular purpose. The specifications in these types of blocks or nets are usually given in terms of absolute tolerances for the coordinates. On the other hand, it is known |1|, |3|, |4| and can also be seen in the appendix of this paper, where the formulae developed in |3|, |4| are summarised, that undetected model errors, i.e. gross errors systematic deformations or deformations of control points, cause distortions in the final coordinates of the network points. These distortions may be larger than the specified tolerances and as such might jeopardize the purpose for which the network has been designed. Taking the above into consideration, the design of a special purpose photogrammetric block, geodetic or combined network is defined as:

> The search for the point configuration and/or the observation scheme such that the undetected model errors will not affect the final coordinates of the network points more than the given tolerance. If by different configurations and/or different observation schemes one could also achieve the specifications, then the design with the smaller risk will be preferred.

In the above definition the concept risk of statistical decision theory is used and in particular the socalled Bayer risk |2|, |7|.

2. The hypothesis testing as a two action decision problem

In statistical decision theory, decisions are also called actions. Actions carry consequences, a way to value the consequences of possible actions, is by considering the loss that could be incurred for each possible action.

The detection of model errors, as can be seen in the appendix, is based on testing the null hypothesis H_o.

$$H_o: E\{\underline{x}^i\} = \hat{x}^i \qquad i, j=1, \ldots, n$$

against the alternative hypothesis H_a

$$H_a: E\{\underline{x}^i\} = \hat{x}^i + C^i_k \nabla^k \qquad k, l=1, \ldots, m$$

$$-\infty < \nabla^k < \infty$$

The hypothesis testing is a two action decision problem. Based on the value of a test variate, which is a function of the observations, one should decide to accept or reject the hypothesis H_o.

A statistical test always implies the risk of taking wrong actions, namely reject H_o even if it is correct or accept H_o even if it is false, the so called type I and type II errors respectively. The probability of committing a type I error is equal to the significance level α of the test, whereas the probability of committing a type II error is $(1-\beta)$, where β is the power of the test. The probabilities α and β can be evaluated from the integrals

$$\alpha = \int_c^{\infty} f\ (x/H_o)\ dx = \lim_{x\to\infty} \{F\ (x/H_o)\} - F\ (c/H_o) \qquad (2.1)$$

$$\beta = \int_c^{\infty} f\ (x/H_a)\ dx = \lim_{x\to\infty} \{F\ (x/H_a)\} - F\ (c/H_a) \qquad (2.2)$$

where $f\ (x/H_o)$ and $f\ (x/H_a)$ are the probability density functions of the test variate under H_o and H_a respectively, and c is the critical value of the test.

Now, assuming that the loss is zero if a correct action is taken during hypothesis testing and is K_i, i=0,1 if an incorrect action is taken, that is to say K_o and K_1 are the losses of type I and type II errors, then it can be proved |4|, |6| that the total risk, or Bayer risk r of taking wrong actions in hypothesis testing is

$$r = \Pi_o K_o \alpha + \Pi_1 K_1\ (1-\beta) \qquad (2.3)$$

where Π_o, $\Pi_1 = 1-\Pi_o$ are the relative occurancies of the hypotheses H_o and H_a based on prior information and α, β, K_o, K_1 as have been defined previously.

In this paper no attempt will be made to quantify the parameters Π_o, K_o, K_1 which are included in (2.3) as in |6|, |5|. The function in (2.3) will be used conceptually in order to define the strategy for a network design based on risk reduction.

3. Design strategy. The target function.

As has been stated in the appendix, model errors of a magnitude

$$\lambda = \frac{1}{\sigma^2}\ \nabla^l g_{lk} \nabla^k \qquad k,\ l=1,\ \ldots,\ m$$

have a probability β to be detected by the test, and thus a probability $(1-\beta)$ to remain undetected. Undetected model errors will affect the final coordinates of the network points. The application of formula (A.14) for the coordinates of the points in a special purpose network, gives an upper bound for this effect

$$\nabla Y^a \leq \sqrt{\bar{\Theta}_o}\ \sigma_{Y^a} \qquad a=1,\ \ldots,\ s \qquad (3.1)$$

where

$$\bar{\Theta}_o = \lambda\ (\mu_i)_{max} \qquad (3.2)$$

as has been defined in (A.12), σ_{Y^a} the standard deviation of the coordinate Y^a and s the number of coordinates.

If a tolerance for the coordinates is specified say, $|\nabla Y|_{tol}$ then, according to the definition in paragraph 1, the network will fulfil the purpose for which it was designed if the relationship

$$\nabla Y^a \leq \sqrt{\bar{\Theta}_o}\ \sigma_{Y^a} = |\nabla Y|_{tol} \qquad a=1, \ldots, s \tag{3.3}$$

is true for every value of a.

The inequalities in (3.3) are satisfied if

$$\sqrt{\bar{\Theta}_o}\ (\sigma_{Y^a})_{max} = |\nabla Y|_{tol} \tag{3.4}$$

where $(\sigma_{Y^a})_{max}$ is the maximum standard deviation of the coordinates of the points.

By making use of (3.2), we obtain

$$\lambda = \frac{(\nabla Y)^2_{tol}}{(\mu_i)_{max}(\sigma_{Y^a Y^a})_{max}} \tag{3.5}$$

where $(\sigma_{Y^a Y^a})_{max}$ is the maximum variance of the coordinates.

Thus, the network satisfies the specifications if the test of model errors expressed by the H_a, detect errors of a magnitude equal to the one given in (3.5) with a probability β. The quantity λ, for a given m (degrees of freedom), is a function of the significance level α and the power of the test β, this means that any value of λ can be achieved for an appropriate pair of values of α and β. Moreover, as can be seen in the charts in |1|, the same value of λ can be obtained from different pairs of values of α and β. From these considerations it is obvious that for a value of λ evaluated from (3.5), the choice of α and β should be restricted. The restriction which is placed here is that the values of α, β should minimize the Bayer risk in (2.3).

From (2.1) and (2.2) it appears that α and β are both functions of the critical value c of the test, thus the Bayer risk as a function of α, β is a function of c.

Minimization of the risk function r is in fact a minimization with respect to c, i.e., we must solve c from the equation.

$$\frac{dr}{dc} = \frac{\partial r}{\partial \alpha} \cdot \frac{d\alpha}{dc} + \frac{\partial r}{\partial \beta}\frac{d\beta}{dc} = 0 \tag{3.6}$$

From (2.3), (2.1) and (2.2) we obtain

$$\frac{\partial r}{\partial \alpha} = \Pi_o K_o, \qquad \frac{\partial r}{\partial \beta} = -\Pi_1 K_1$$
$$\frac{d\alpha}{dc} = -\frac{dF(c/Ho)}{dc} = -f\,(c/H_o)$$
$$\frac{d\beta}{dc} = -\frac{dF(c/Ha)}{dc} = -f\,(c/H_a)$$

The equation (3.6) then becomes

$$-\Pi_o K_o f(c/H_o) + \Pi_1 K_1 f(c/H_a) = 0 \qquad (3.7)$$

The solution of (3.7) with respect to c introduced in (2.1) and (2.2) will give the values of α, β, say α_o and β_o which minimize the risk function, thus the minimum risk is

$$r_{min} = \Pi_o K_o \alpha_o + \Pi_1 K_1 (1-\beta_o) \qquad (3.8)$$

For any new network design the denominator in (3.5) will change and thus λ will take other value, which consequently, will give different minimum risk. Each network design is characterized by its minimum risk. Among the different networks, the one with the smaller minimum risk should be preferred. It is therefore evident that the r_{min} in (3.8) plays here the role of the target function whose minimization is attempted by changing the network design. As has been stated in paragraph 2, no attempt will be made to quantify the parameters Π_o, K_o, K_1, so that it is not possible to evaluate the r_{min}. It will then be replaced by another quantity which can readily be used for choosing a better, with respect to risk, network.

The risk function in (2.3) depends on the parameter λ, and can be written symbolically as

$$r = r(\alpha, \beta) \qquad (3.9)$$

where α and β are functions of the critical value c and λ

$$\alpha = \alpha(c), \quad \beta = \beta(c, \lambda)$$

and c itself is a function of λ

$$c = c(\lambda)$$

The total derivative of (3.9) with respect to λ is

$$\frac{dr}{d\lambda} = \frac{\partial r}{\partial \alpha}\frac{\partial \alpha}{\partial c}\frac{dc}{d\lambda} + \frac{\partial r}{\partial \beta}\frac{\partial \beta}{\partial c}\frac{dc}{d\lambda} + \frac{\partial r}{\partial \beta}\frac{\partial \beta}{\partial \lambda}$$

or

$$\frac{dr}{d\lambda} = \left(\frac{\partial r}{\partial \alpha}\frac{\partial \alpha}{\partial c} + \frac{\partial r}{\partial \beta}\frac{\partial \beta}{\partial c}\right)\frac{dc}{d\lambda} + \frac{\partial r}{\partial \beta}\frac{\partial \beta}{\partial \lambda}$$

The terms inside the parenthesis is the derivative of r with respect to c, so the above relationship can be written

$$\frac{dr}{d\lambda} = \frac{\partial r}{\partial c}\frac{dc}{d\lambda} + \frac{\partial r}{\partial \beta}\frac{\partial \beta}{\partial \lambda} \qquad (3.10)$$

for the minimum risk, we have

$$\frac{\partial r}{\partial c} = 0$$

using the above relation, (3.10) becomes

$$\frac{dr}{\partial \lambda} = \frac{\partial r}{\partial \beta}\frac{\partial \beta}{\partial \lambda} \qquad (3.11)$$

the evaluation of $\frac{\partial r}{\partial \beta}$ gives

$$\frac{\partial r}{\partial \beta} = -\Pi_1 K_1$$

as far as the second term in (3.11) is concerned, it has been proved in |4| that it is equal to

$$\frac{\partial \beta}{\partial \lambda} = \frac{e^{-\lambda/2}}{m} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\lambda}{2}\right)^n G\left(c/\frac{m}{2},\ n+1+\frac{m}{2}\right) \qquad (3.12)$$

where the symbol G (./.,.) which is defined as

$$G(x/a,p) = \frac{a^{\rho}}{\Gamma(\rho)} e^{-ax} x^{\rho-1} \quad \text{for } a > 0,\ \rho > 0,\ 0 < x < \infty$$

is the probability density function of the Gamma distribution |8|. It is observed that the relation in (3.12) is always positive. Now, the (3.11) can be written as

$$\frac{\partial r}{\partial \lambda} = -\Pi_1 K_1 \frac{e^{-\lambda/2}}{m} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\lambda}{2}\right)^n G\left(c/\frac{m}{2},\ n+1+\frac{m}{2}\right)$$

due to the minus sign, it is concluded that

$$\frac{\partial r}{\partial \lambda} < 0$$

which indicates that the minimum risk is a decreasing function of λ. That means that among the different networks, the one with the greater λ calculated from (3.5) should be preferred, since it will have a smaller minimum risk. The numerator in (3.5) has a fixed value, expressing the specifications. The value of λ depends upon the denominator. A greater λ is achieved as the denominator becomes smaller. Hence, the criterion for a better network is based on the value of the quantity.

$$M = (\mu_i)_{max} (\sigma_Y a_Y a)_{max} \qquad (3.13)$$

which is proposed to be used as a target function for a network design based on risk reduction. The smaller the value of M, meaning smaller risk, the better the network.

4. Epilogue

The choice of the upper bound $\bar{\Theta}_o$ in paragraph 3 and consequently the use of $(\mu_i)_{max}$ in (3.5), indicates that per network, say i, the minimum magnitude of the model errors to be detected is considered. Symbolically this can be indicated as

$$\text{network } i \dashrightarrow \lambda^i_{min} \qquad (4.1)$$

If instead of the upper bound $\bar{\Theta}_o$, another value was selected for the effect of the undetected errors on the coordinates, then the magnitude of the model errors to be detected, will be greater then λ^i_{min}. Now, taking into account that the minimum risk is a decreasing function of λ it can be concluded that the r_{min} in (3.8) is in fact the maximum of the minimum risks and this is the one which characterizes each network. Symbolically we can write

$$\text{network } i \dashrightarrow r^i_{maxmin} = r^i_{min}(\lambda^i_{min}) \qquad (4.2)$$

Furthermore, the criterion for a better network, which as we have seen, is based on the reduction of the value of the quantity M in (3.13), means in

terms of the risk function, that by changing the network design a minimization of the r^i_{maxmin} over the different networks i is attempted. Thus, we have

$$(r^i_{maxmin})_{min} = \min_i (r^i_{maxmin}) \tag{4.3}$$

The steps expressed in (4.1) to (4.3) outline conceptually the design strategy for special purpose networks by risk reduction.

The target function M in (3.13) depends upon the precision $(\sigma_Y a_Y a)_{max}$ of the coordinates of the network points and upon the parameter $(\mu_i)_{max}$ which is involved in the definition of the external reliability (see appendix). Hence, the design of a reduced risk network, takes into account the two parameters of the quality control of a point determination system, namely precision and reliability.

Appendix

Model errors, i.e. gross errors, systematic deformations or deformations of terrain points can be expressed as alternative hypotheses in the form

$$H_a: E\{\underline{x}^i\} = \hat{x}^i + c^i_k \nabla^k \qquad \begin{matrix} i, j=1, \ldots, n \\ k, l=1, \ldots, m \end{matrix} \tag{A.1}$$
$$-\infty < \nabla^k < \infty$$

where, $\underline{x}^i$ are the observations and $\hat{x}^i$ their mean values.

The possible presence of model errors can then be tested by testing the null hypothesis H_o.

$$H_o: E\{\underline{x}^i\} = \hat{x}^i \tag{A.2}$$

against the alternative hypothesis H_a. The test is based on the test quantity

$$\underline{\omega}_g = \frac{\underline{\omega}_k g^{kl} \underline{\omega}_l}{m\sigma^2} \tag{A.3}$$

with

$$\underline{\omega}_l = c^j_l \, g_{ji} \, (-\Delta x^i) \tag{A.4}$$

and

$$g_{lk} = c^j_l \, g_{ji} (g^{ij} - G^{ij}) g_{ji} \, c^i_k \tag{A.5}$$

where, Δx^i is the least square corrections to the observations
g^{ij} the weight coefficient matrix of the observations
G^{ij} the weight coefficient matrix of the corrected observations
g_{ji} is the inverse of g^{ij}
g^{kl} is the inverse of g_{lk}

The probability density function of the test quantity $\underline{\omega}_g$ in (A.3) under H_o and H_a are respectively

$$f(x/H_o) = G(x/\tfrac{m}{2}, \tfrac{m}{2})$$

$$f(x/H_a) = e^{-\lambda/2} \sum_{n=0}^{\infty} \frac{1}{n!} (\tfrac{\lambda}{2})^n \, G(x/\tfrac{m}{2}, n + \tfrac{m}{2}) \qquad (A.6)$$

where, the symbol G(./.,.) defined as

$$G(x/\alpha,p) = \frac{\alpha^\rho}{\Gamma(\rho)} e^{-\alpha x} x^{\rho-1} \qquad \text{for } \alpha > 0,\ \rho > 0,\ 0 < x < \infty$$

is the probability density function of the Gamma distribution, and λ, the eccentricity parameter

$$\lambda = \frac{1}{\sigma^2} \nabla^l g_{lk} \nabla^k \qquad (A.7)$$

σ^2 is the variance factor.

A choice of a significance level α and the corresponding critical value c leads to

$$\text{reject } H_o \text{ if } \underline{\omega}_g > c$$
$$\text{accept } H_o \text{ if } \underline{\omega}_g \leq c \qquad (A.8)$$

The test in (A.8) detects errors of a magnitude

$$\lambda = \frac{1}{\sigma^2} \nabla^l g_{lk} \nabla^k \qquad (A.9)$$

with a probability β (the power of the test). Thus, there is a probability $(1-\beta)$ that errors of a magnitude given in (A.9) will remain undetected.

Undetected model errors will affect the unknown parameters in an adjustment problem. The square norm of the vector of this effect, is given by

$$\Theta_o = \frac{1}{\sigma^2} \nabla^l \underline{g}_{lk} \nabla^k \qquad (A.10)$$

with

$$\underline{g}_{lk} = c^j_l g_{ji} \, G^{ij} g_{ji} \, c^i_k \qquad (A.11)$$

An upper bound of Θ_o is

$$\bar{\Theta}_o = \lambda \, (\mu_i)_{max} \qquad (A.12)$$

where $(\mu_i)_{max}$ is the maximum eigen value of the eigen value problem.

$$\underline{g}_{lk} \, v_i = \mu_i g_{lk} v_i \qquad (A.13)$$

The upper bound given in (A.12) is a measure of assessing the external relability for a point determination system, due to the fact that if $f = f(\ldots Y^a \ldots)$ is an arbitrary function of the parameters, i.e., coordinates of the points in the case of a point determination problem and if ∇f represents the effect of undetected model errors on this function, then

$$|\nabla f| \leq \sqrt{\bar{\Theta}_o} \cdot \sigma_f \qquad (A.14)$$

where σ_f is the standard deviation of the function f.
That is to say, the effect of undetected model errors on the function f, cannot be larger than $\sqrt{\bar{\Theta}_o}$ times the standard deviation of the function.

References

1. Baarda, W. A testing procedure for use in Geodetic networks. Netherlands Geodetic Commission, 1968.

2. Berger, J.O. Statistical Decision Theory. Springer Verlag, New York, Heidelberg, Berlin, 1980.

3. Bouloucos, T. Multidimensional tests for model errors and their reliability measures. Proc. Comm. III ISPRS, Rovaniemi Finland 1986, ITC Journal 1986-3.

4. Bouloucos, T. A unified approach for the detection of model errors. Technical University of Athens, 1988.

5. Mierlo, J. van Problems of computing costs in decision problems, Stuttgart, Photogrammetric Week, 1983.

6. Molenaar, M. Risk minimization for error detection procedures in Photogrammetry. ITC Journal 1981-2.

7. Neveu, J. Mathematical foundation of the calculus of probability. Holden-Day Inc., San Fransisco, London, Amsterdam, 1965.

8. Rao, C.R. Linear statistical inference and its applications. J. Wiley and Sons, 1973.