

# Optimal distances

- Study a linear model, e.g.

$$\mathbf{g}(t) = \mathbf{p} + t\mathbf{v},$$

with jacobian  $J(t) = \mathbf{v}$ .

- Assume we want to estimate the model parameters that best describe our measurements  $\mathbf{b}$ .
- If the measurement errors in  $\mathbf{b}$  are  $N(0, \sigma^2 \mathbf{I})$ , i.e. have a covariance matrix  $\Sigma = \sigma^2 \mathbf{I}$ , then the optimal solution is

$$\hat{t} = (\mathbf{v}^\top \mathbf{v})^{-1} \mathbf{v}^\top (\mathbf{b} - \mathbf{p})$$

that corresponds to the solution of

$$\min_t \|\mathbf{g}(t) - \mathbf{b}\|^2 = (\mathbf{g}(t) - \mathbf{b})^\top \mathbf{I}^{-1} (\mathbf{g}(t) - \mathbf{b})$$

- p. 1

# Optimal distances

- In the general case with covariance matrix  $\Sigma$  the solution instead becomes

$$\hat{t} = (\mathbf{v}^\top \Sigma^{-1} \mathbf{v})^{-1} \mathbf{v}^\top \Sigma^{-1} (\mathbf{b} - \mathbf{p})$$

that corresponds to the solution of

$$\min_t \|\mathbf{g}(t) - \mathbf{b}\|_\Sigma^2 = (\mathbf{g}(t) - \mathbf{b})^\top \Sigma^{-1} (\mathbf{g}(t) - \mathbf{b}),$$

where  $\Sigma^{-1}$  enters as a weight matrix.

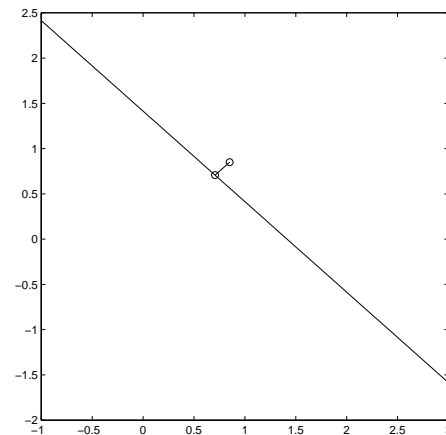
- p. 3

# Optimal distances, example

$$\mathbf{p} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

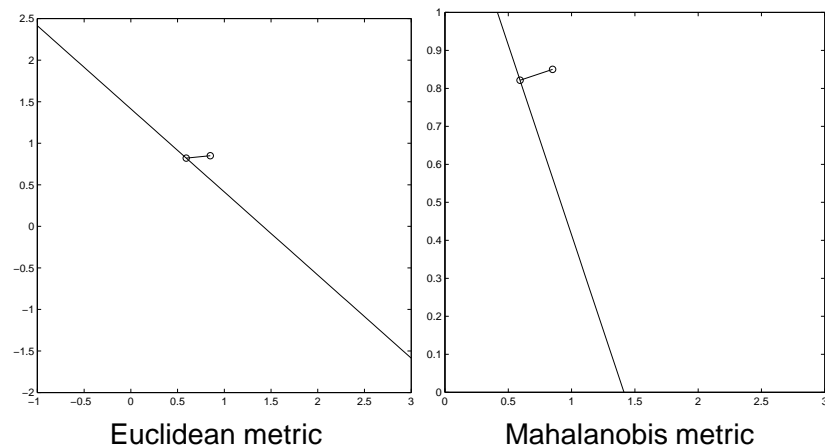
$$\mathbf{b} = \begin{bmatrix} 0.85 \\ 0.85 \end{bmatrix}, \Sigma = 10^{-4} \mathbf{I},$$

$$\hat{t} = 0.71, \mathbf{g}(\hat{t}) = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}$$



# Optimal distances, example

$$\mathbf{p} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0.85 \\ 0.85 \end{bmatrix}, \Sigma = 10^{-4} \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}, \hat{t} = 0.59, \mathbf{g}(\hat{t}) = \begin{bmatrix} 0.59 \\ 0.82 \end{bmatrix}$$



Euclidean metric

Mahalanobis metric

# The variance of the solution

The covariance of the solution  $\hat{t}$  becomes

$$\Sigma_{\hat{t}} = (\mathbf{v}^\top \Sigma^{-1} \mathbf{v})^{-1}$$

and for  $\hat{\mathbf{g}} = \mathbf{g}(\hat{t})$

$$\Sigma_{\hat{\mathbf{g}}} = \mathbf{v} \Sigma_{\hat{t}} \mathbf{v}^\top = \mathbf{v} (\mathbf{v}^\top \Sigma^{-1} \mathbf{v})^{-1} \mathbf{v}^\top$$

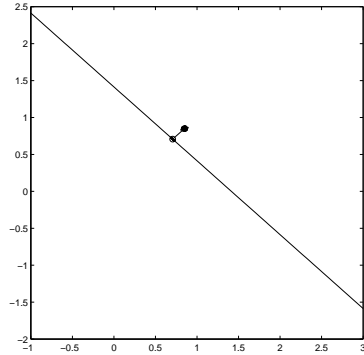
$$\Sigma = 10^{-4} \mathbf{I}, \quad \hat{t} = 0.71,$$

$$\Sigma_{\hat{t}} = 0.5 \cdot 10^{-4}, \quad \widehat{\Sigma}_{\hat{t}} = 0.51 \cdot 10^{-4},$$

$$\hat{\mathbf{g}} = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix},$$

$$\Sigma_{\hat{\mathbf{g}}} = 0.5 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\widehat{\Sigma}_{\hat{\mathbf{g}}} = 0.51 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$



- p. 5

# Difference between unweighted and weighted

If the measurement error is anisotropic, the unweighted solution will have a larger variance than the weighted:

Unweighted solution:

Weighted solution:

$$\Sigma = 10^{-4} \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\Sigma = 10^{-4} \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\hat{t} = 0.71,$$

$$\hat{t} = 0.59,$$

$$\widehat{\Sigma}_{\hat{t}} = 2.6 \cdot 10^{-4},$$

$$\widehat{\Sigma}_{\hat{t}} = 0.93 \cdot 10^{-4},$$

$$\hat{\mathbf{g}} = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix},$$

$$\hat{\mathbf{g}} = \begin{bmatrix} 0.59 \\ 0.82 \end{bmatrix},$$

$$\widehat{\Sigma}_{\hat{\mathbf{g}}} = 2.6 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

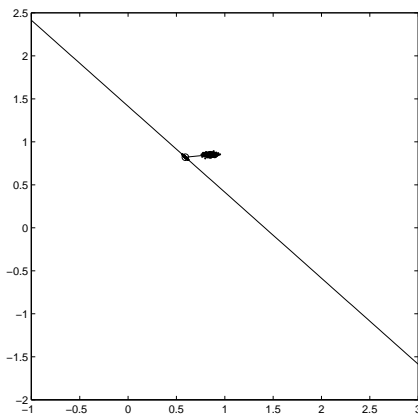
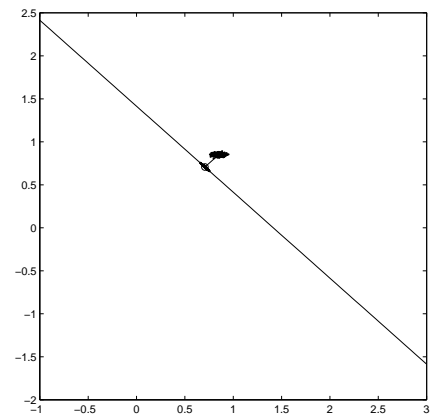
$$\widehat{\Sigma}_{\hat{\mathbf{g}}} = 0.93 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

# Difference between unweighted and weighted

Unweighted solution:

Weighted solution:

$$\hat{\mathbf{g}} = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}, \widehat{\Sigma}_{\hat{\mathbf{g}}} = 2.6 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \hat{\mathbf{g}} = \begin{bmatrix} 0.59 \\ 0.82 \end{bmatrix}, \widehat{\Sigma}_{\hat{\mathbf{g}}} = 0.93 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$



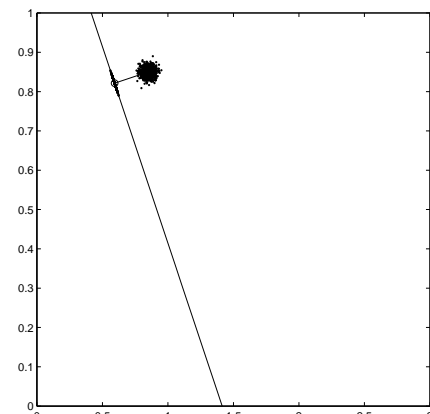
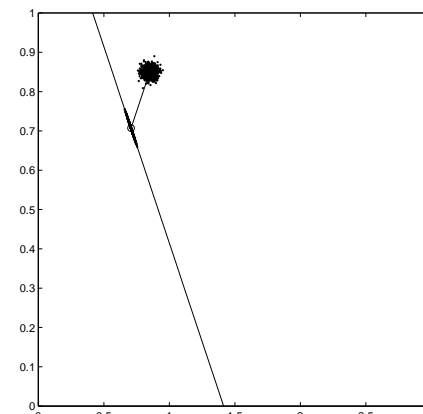
- p. 7

# Same thing in Mahalanobis space

Unweighted solution:

Weighted solution:

$$\hat{\mathbf{g}} = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}, \widehat{\Sigma}_{\hat{\mathbf{g}}} = 2.6 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \hat{\mathbf{g}} = \begin{bmatrix} 0.59 \\ 0.82 \end{bmatrix}, \widehat{\Sigma}_{\hat{\mathbf{g}}} = 0.93 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$



# Constraint formulation

- A constrained formulation of the same problem

$$\min_{\mathbf{x}} \|\mathbf{g}(\mathbf{x}) - \mathbf{b}\|_{\Sigma}^2$$

$$\text{s.t. } \mathbf{c}(\mathbf{x}) = \mathbf{0},$$

where  $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ ,  $\mathbf{J}(\mathbf{x}) = \mathbf{I}$  and  $\mathbf{c}(\mathbf{x}) = \mathbf{w}^T(\mathbf{x} - \mathbf{p})$ , where  $\mathbf{w}$  is a basis for the nullspace of the constraint  $\mathbf{w}^T \mathbf{v} = 0$ .

- In general with  $\mathbf{A}$  as jacobian to the constraint the covariance matrix for  $\hat{\mathbf{g}}$  becomes

$$\Sigma_{\hat{\mathbf{g}}} = \mathbf{A}(\mathbf{A}^T \mathbf{J}^T \Sigma^{-1} \mathbf{J} \mathbf{A})^{-1} \mathbf{A}^T.$$

## Confidence limits

- Given the covariance matrix  $\Sigma_{\hat{\mathbf{g}}}$  it is possible to construct confidence limits for  $\hat{\mathbf{g}}$  such that

$$\mathbf{x}^T \Sigma_{\hat{\mathbf{g}}}^{-1} \mathbf{x} = k^2,$$

where  $k^2 = F_n^{-1}(\alpha)$  from the  $\chi^2$  distribution with  $n$  degrees of freedom and at the confidence level  $\alpha$ .

- With the eigenvalue factorization  $\mathbf{V} \mathbf{D} \mathbf{V}^T = \Sigma_{\hat{\mathbf{g}}}$  we may determine confidence ellipsoids from

$$\mathbf{x}^T \Sigma_{\hat{\mathbf{g}}}^{-1} \mathbf{x} = k^2,$$

$$\mathbf{x}^T \mathbf{V} \mathbf{D}^{-1} \mathbf{V}^T \mathbf{x} = k^2,$$

where the columns of  $\mathbf{V}$  are the direction of the main axes of the ellipsoid and  $1/\sqrt{d_{ii}}$  are the lengths of the main axes.

# Non-linear models

- För non-linear models the covariance formulas are *first order approximations*.

$$\Sigma = 10^{-4} \mathbf{I},$$

$$\hat{t} = 0.71,$$

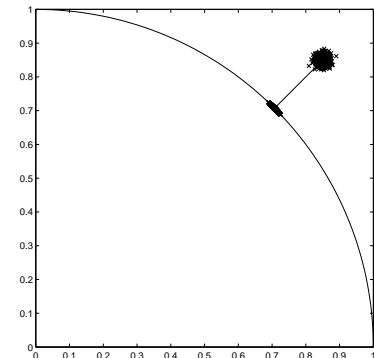
$$\Sigma_{\hat{t}} = 1 \cdot 10^{-4},$$

$$\widehat{\Sigma}_{\hat{t}} = 0.71 \cdot 10^{-4},$$

$$\hat{\mathbf{g}} = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix},$$

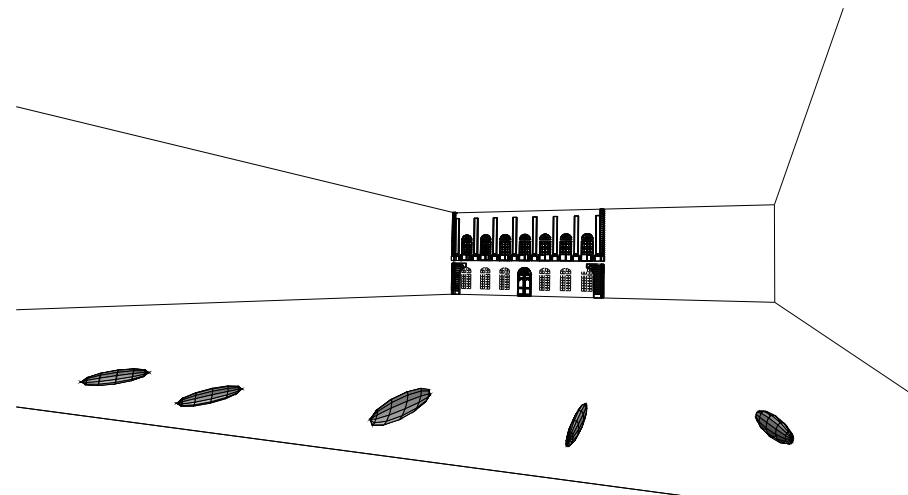
$$\Sigma_{\hat{\mathbf{g}}} = 0.5 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\widehat{\Sigma}_{\hat{\mathbf{g}}} = 0.36 \cdot 10^{-4} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$



- p. 9

## Example



- p. 11