SET MEMBERSHIP ESTIMATION THEORY

Michele TARAGNA

Dipartimento di Elettronica e Telecomunicazioni

Politecnico di Torino

michele.taragna@polito.it

Master Course in Mechatronic Engineering

Master Course in Computer Engineering

01RKYQW / 01RKYOV “Estimation, Filtering and System Identification”

Academic Year 2019/2020

Example: estimation of a resistance value

N voltage-current measurements are performed on a real resistor, assuming that:

- its static characteristic is linear $\Rightarrow$ the device model is given by the Ohm's law

$$v_R = R \cdot i_R$$

- the measurements are corrupted by an unknown noise

$$e = [e_1, \ldots, e_N]^T$$

The following system of linear equations is derived:

$$\begin{cases} v_{R,1} & = & R \cdot i_{R,1} + e_1 \\ v_{R,2} & = & R \cdot i_{R,2} + e_2 \\ & \vdots & \\ v_{R,N} & = & R \cdot i_{R,N} + e_N \end{cases}$$

In matrix terms:

$$\underbrace{\begin{bmatrix} v_{R,1} \\ v_{R,2} \\ \vdots \\ v_{R,N} \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} i_{R,1} \\ i_{R,2} \\ \vdots \\ i_{R,N} \end{bmatrix}}_{\Phi} \cdot \underbrace{[R\,]}_{\theta_o} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}}_{e}$$

is in the standard form:

$$\underbrace{\mathbf{y}}_{\text{known data}} = \underbrace{\mathbf{F}(\theta_o)}_{\text{known function}} + \underbrace{\mathbf{e}}_{\text{unknown noise}}$$

$$F(\theta_o) = \Phi \cdot \theta_o = \text{ linear function of the unknown parameter } \theta_o$$

Goal: find an estimate $\hat{R}$ of R by means of an estimation algorithm (estimator) ψ applied to the data vector y:

$$\hat{R} = \psi(y) \cong R$$

Least squares estimation errors

θ_o : "true" parameters that generated the data vector y

Due to measurement noise, $y = \Phi\theta_o + e \neq \Phi\theta_o \quad \Rightarrow$

using the least squares algorithm as estimator:

$$\begin{aligned}\hat{\theta} &= \left(\Phi^T\Phi\right)^{-1}\Phi^T y = \left(\Phi^T\Phi\right)^{-1}\Phi^T\left(\Phi\theta_o + e\right) = \\ &= \underbrace{\left(\Phi^T\Phi\right)^{-1}\Phi^T\Phi}_{I}\,\theta_o + \left(\Phi^T\Phi\right)^{-1}\Phi^T e = \theta_o + \left(\Phi^T\Phi\right)^{-1}\Phi^T e\end{aligned}$$

$$\Downarrow$$

$$\hat{\theta} - \theta_o = \left(\Phi^T\Phi\right)^{-1}\Phi^T e = \text{ estimation error}$$

- e is not exactly known, but different assumptions may be made on e :
  - random variable $\longrightarrow$ statistical estimation
  - componentwise bounded
  - energy bounded

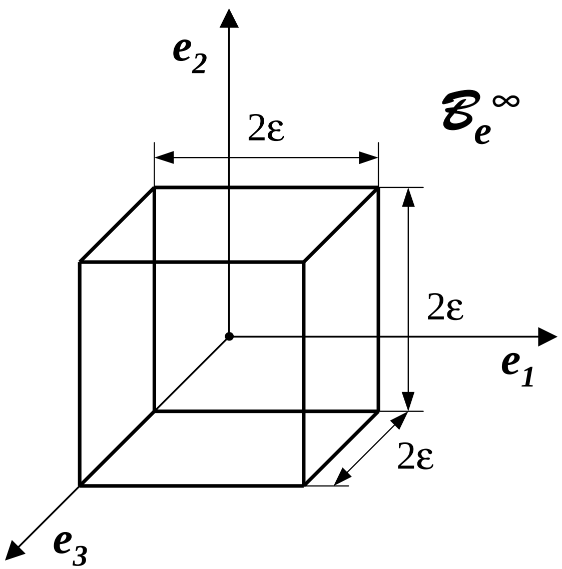
  (componentwise bounded, energy bounded) $\Big\}\longrightarrow$ **Set Membership estimation**

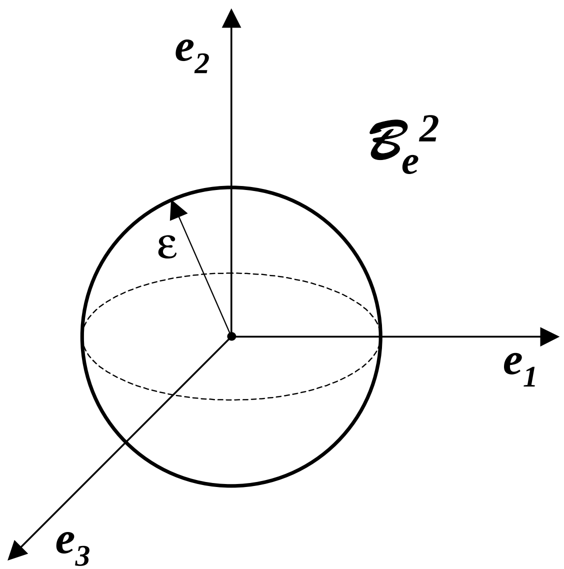
Unknown But Bounded (UBB) errors

$$e \in \mathcal{B}_e = \text{ uncertainty set}$$

$$\mathcal{B}_e^{\infty} = \left\{ \tilde{e} \in \mathbb{R}^N : |\tilde{e}_i| \leq \varepsilon, i = 1, \ldots, N \right\} = \left\{ \tilde{e} \in \mathbb{R}^N : \|\tilde{e}\|_{\infty} = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon \right\}$$

$$\mathcal{B}_e^2 = \left\{ \tilde{e} \in \mathbb{R}^N : \tilde{e}^T \cdot \tilde{e} = \sum_{i=1}^{N} \tilde{e}_i^2 \leq \varepsilon^2 \right\} = \left\{ \tilde{e} \in \mathbb{R}^N : \|\tilde{e}\|_2 = \sqrt{\sum_{i=1}^{N} \tilde{e}_i^2} \leq \varepsilon \right\}$$





- **Assumption**: $\mathcal{B}_e$ is symmetric with respect to the origin of $\mathbb{R}^N$

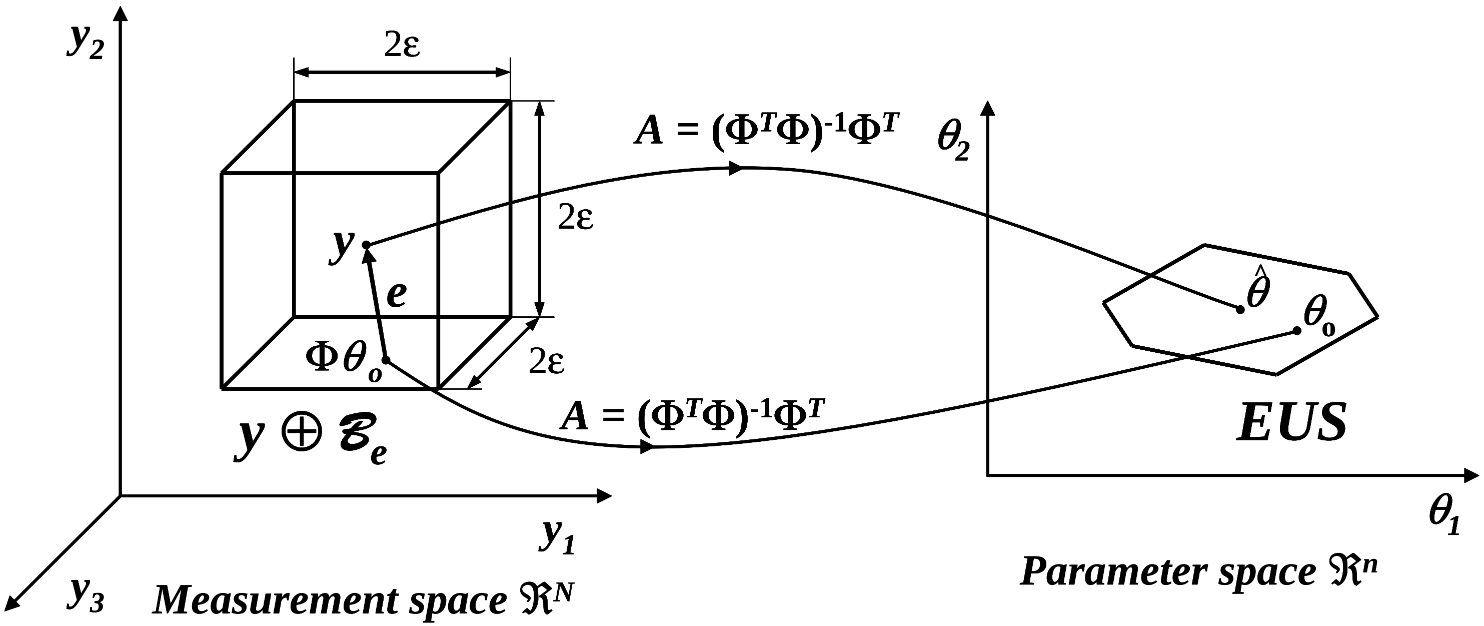
- **Problem**: how to evaluate the uncertainty on $\hat{\theta}$ induced by the uncertainty set $\mathcal{B}_e$?

$$A = \left(\Phi^T \Phi\right)^{-1} \Phi^T = \text{least squares operator : } \underbrace{\mathbb{R}^N}_{\text{measurement space}} \rightarrow \underbrace{\mathbb{R}^n}_{\text{parameter space}}$$

$$\hat{\theta} - \theta_o = \left(\Phi^T \Phi\right)^{-1} \Phi^T e = Ae \;\Rightarrow$$

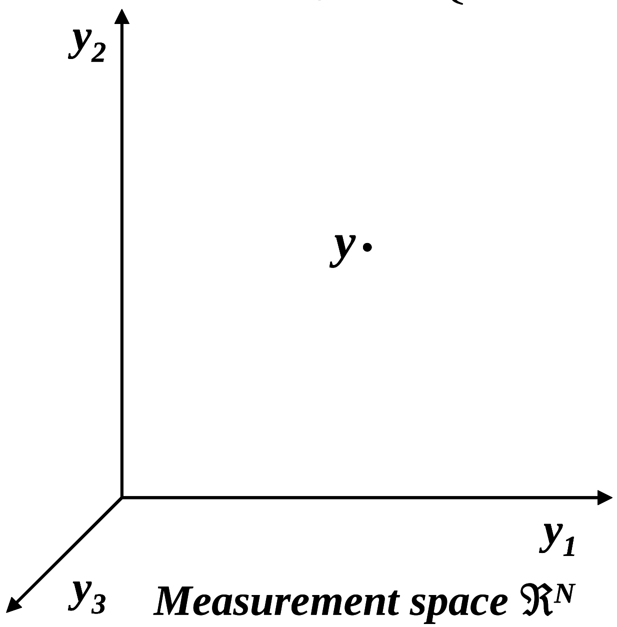
$$\theta_o = \hat{\theta} - Ae \;\Rightarrow$$

$$\theta_o \in EUS = \hat{\theta} \oplus A[\mathcal{B}_e] = Ay \oplus A[\mathcal{B}_e] = A\left[y \oplus \mathcal{B}_e\right] = \textbf{Estimate Uncertainty Set}$$



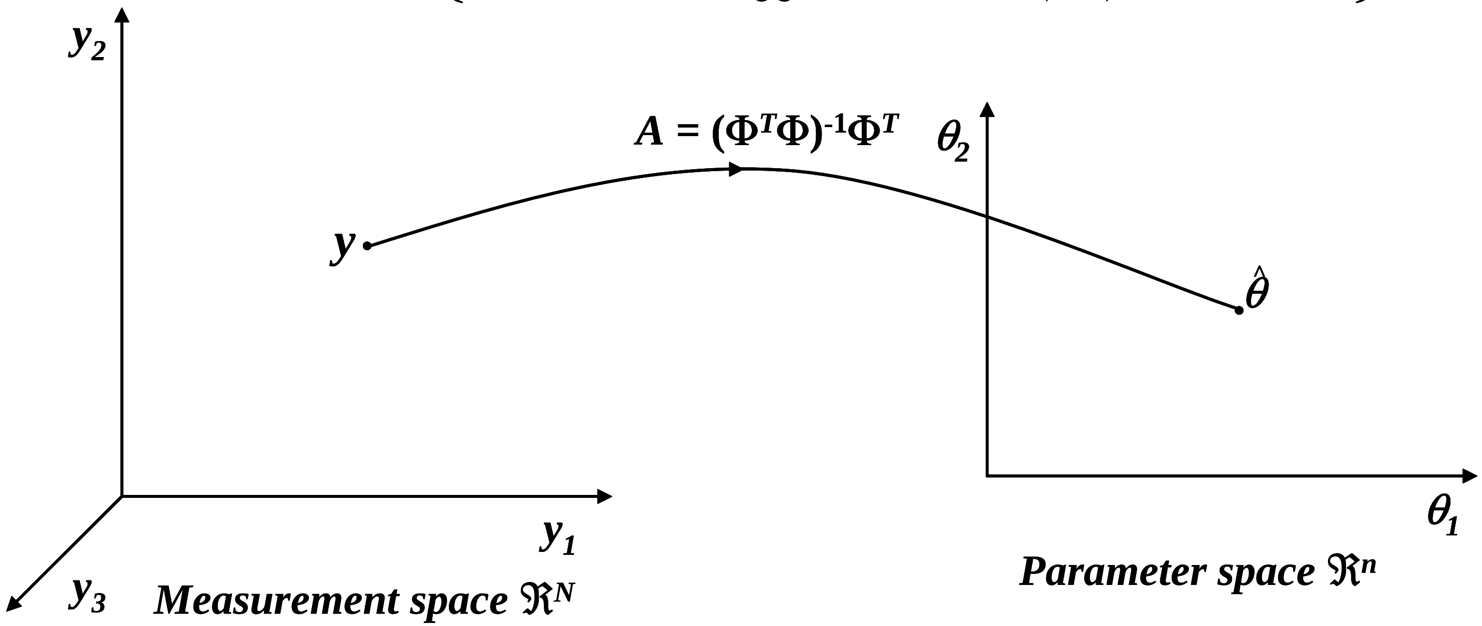
Note that $\theta_o \in EUS$ and that the distance between $\Phi\theta_o$ and y is not greater than ε

Example: let N = number of measurements $= 3$, n = number of parameters $= 2$ and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N \colon \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \le \varepsilon\right\}$ with known ε



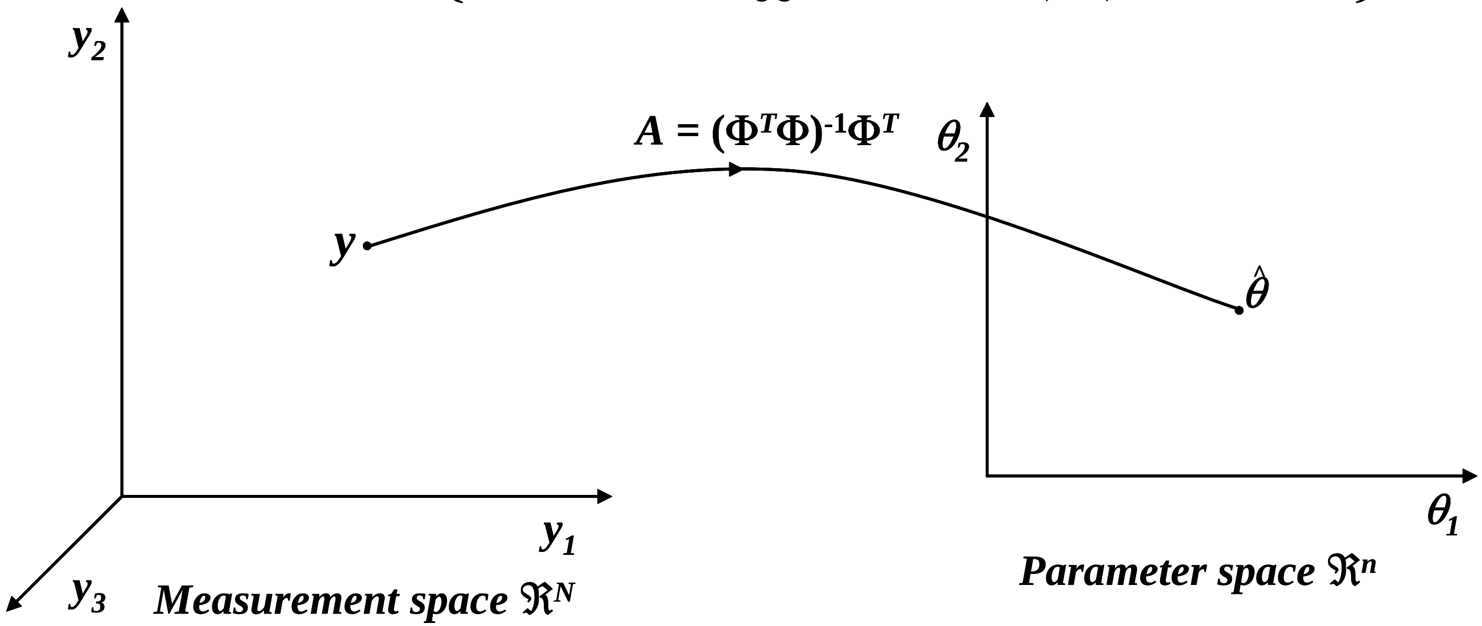
- Each data vector y is an element of the **Measurements space** $\mathbb{R}^N = \mathbb{R}^3$

Example: let N = number of measurements = 3, n = number of parameters = 2 and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N\text{: } \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon\right\}$ with known ε



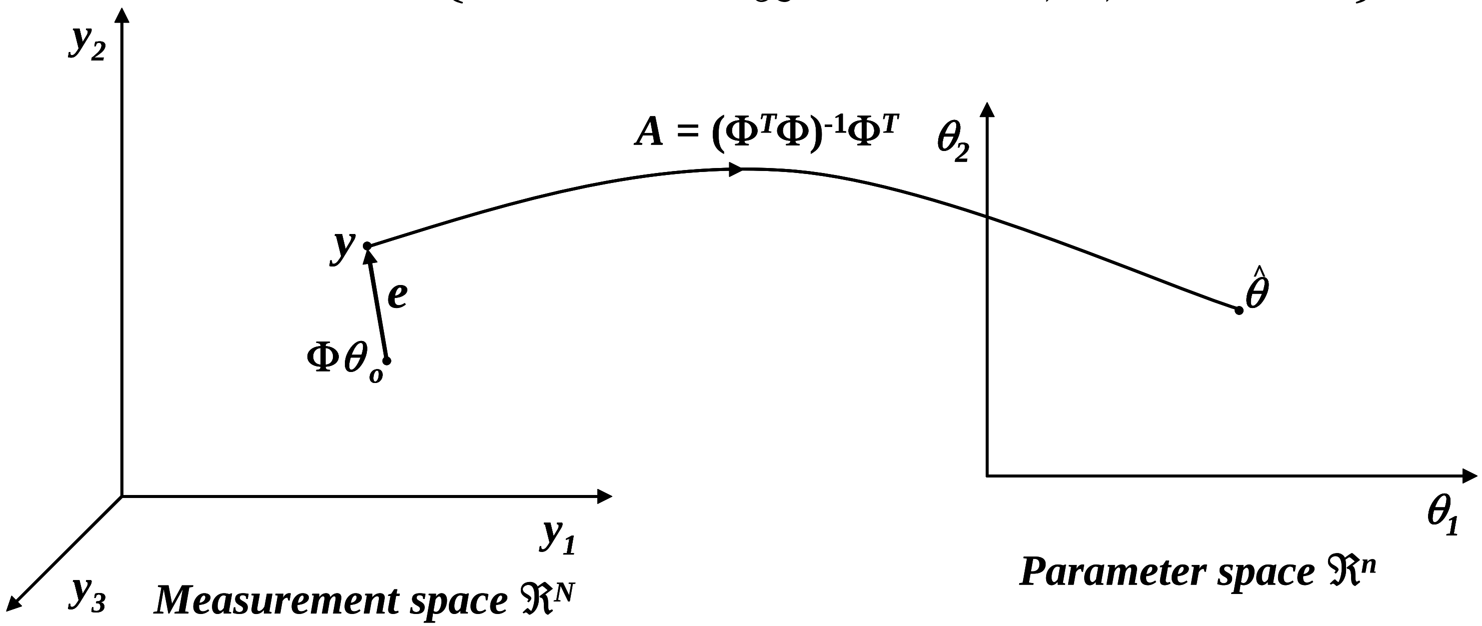
- Each data vector y is an element of the **Measurements space** $\mathbb{R}^N = \mathbb{R}^3$
- If the least squares operator $A = \left(\Phi^T\Phi\right)^{-1}\Phi^T$ is applied as estimator ψ to the data vector y, then the estimate $\hat{\theta} = \psi\left(y\right) = Ay$ is an element of the **Parameter space** $\mathbb{R}^n = \mathbb{R}^2$ obtained as linear mapping of $y \in \mathbb{R}^N$ into $\mathbb{R}^n$

Example: let N = number of measurements = 3, n = number of parameters = 2 and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N : \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon\right\}$ with known ε



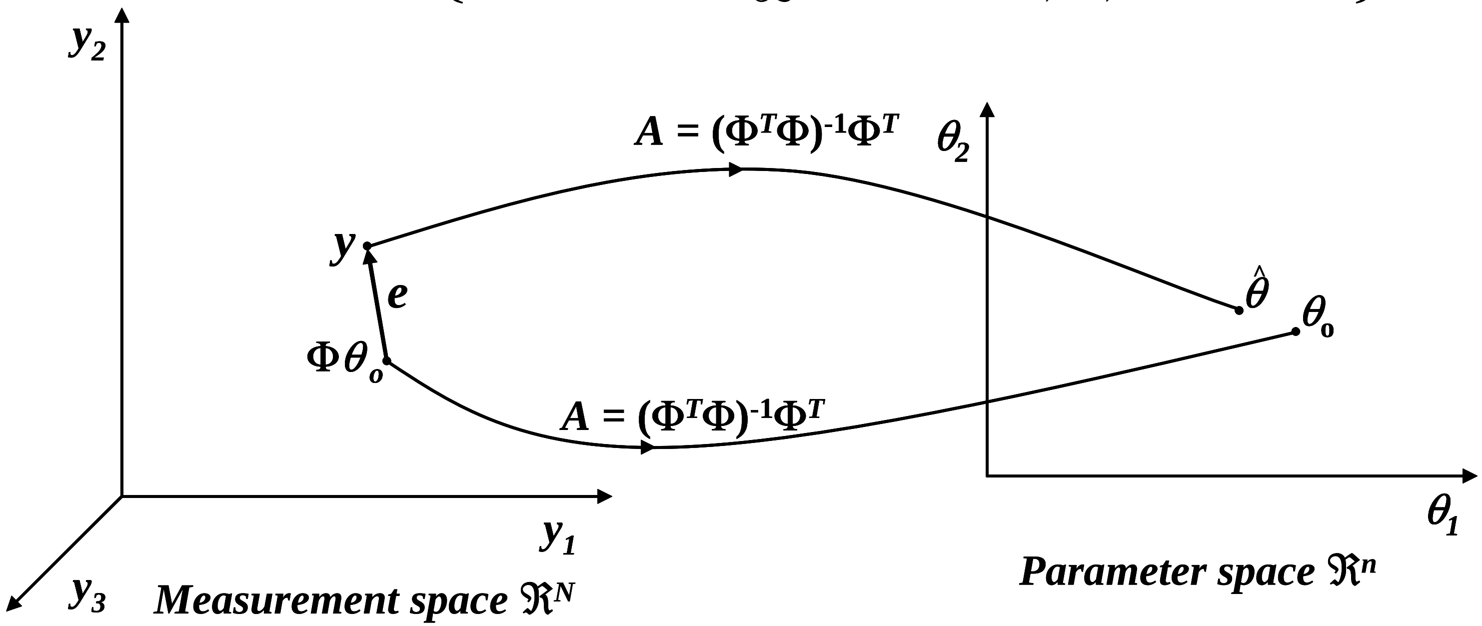
- Each data vector y is an element of the **Measurements space** $\mathbb{R}^N = \mathbb{R}^3$
- If the least squares operator $A = \left(\Phi^T\Phi\right)^{-1}\Phi^T$ is applied as estimator ψ to the data vector y, then the estimate $\hat{\theta} = \psi\left(y\right) = Ay$ is an element of the **Parameter space** $\mathbb{R}^n = \mathbb{R}^2$ obtained as linear mapping of $y \in \mathbb{R}^N$ into $\mathbb{R}^n$
- If $e = \mathbf{0} \Rightarrow \psi\left(y \equiv \Phi\theta_o\right) = A\Phi\theta_o = \left(\Phi^T\Phi\right)^{-1}\Phi^T\Phi\theta_o = \theta_o$

Example: let N = number of measurements $= 3$, n = number of parameters $= 2$ and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N\text{: } \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon\right\}$ with known ε



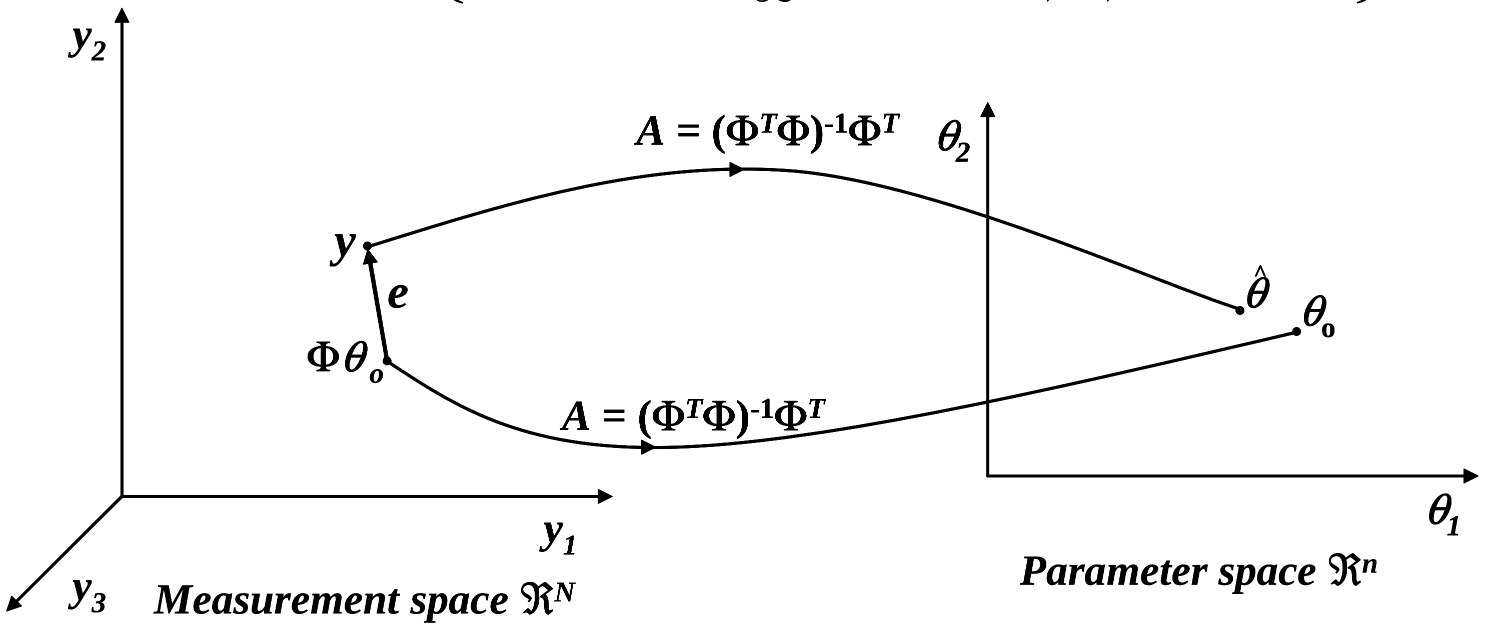
- Each data vector y is an element of the **Measurements space** $\mathbb{R}^N = \mathbb{R}^3$
- If the least squares operator $A = \left(\Phi^T\Phi\right)^{-1}\Phi^T$ is applied as estimator ψ to the data vector y, then the estimate $\hat{\theta} = \psi\,(y) = Ay$ is an element of the **Parameter space** $\mathbb{R}^n = \mathbb{R}^2$ obtained as linear mapping of $y \in \mathbb{R}^N$ into $\mathbb{R}^n$
- If $e = \mathbf{0} \Rightarrow \psi\,(y \equiv \Phi\theta_o) = A\Phi\theta_o = \left(\Phi^T\Phi\right)^{-1}\Phi^T\Phi\theta_o = \theta_o$
- In general, $e \neq \mathbf{0} \Rightarrow y = \Phi\theta_o + e \neq \Phi\theta_o$

Example: let N = number of measurements $= 3$, n = number of parameters $= 2$ and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N\colon \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon\right\}$ with known ε



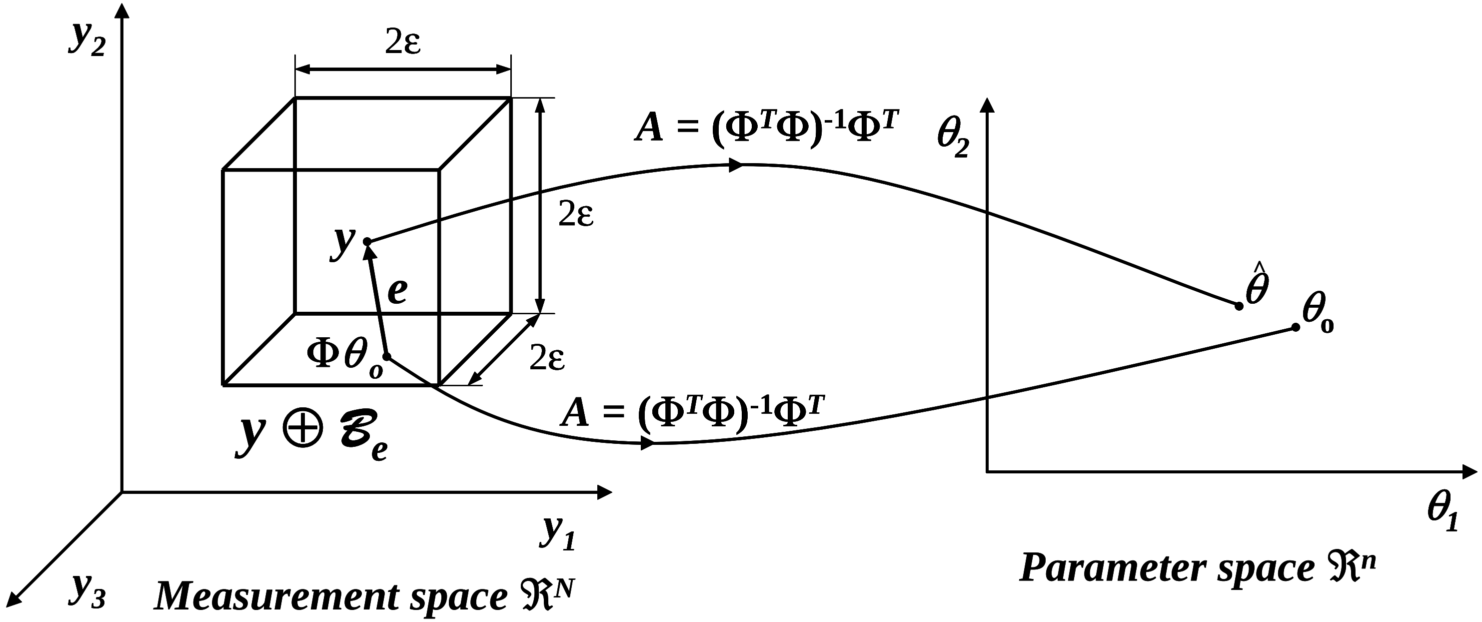
- Each data vector y is an element of the **Measurements space** $\mathbb{R}^N = \mathbb{R}^3$
- If the least squares operator $A = \left(\Phi^T\Phi\right)^{-1}\Phi^T$ is applied as estimator ψ to the data vector y, then the estimate $\hat{\theta} = \psi(y) = Ay$ is an element of the **Parameter space** $\mathbb{R}^n = \mathbb{R}^2$ obtained as linear mapping of $y \in \mathbb{R}^N$ into $\mathbb{R}^n$
- If $e = \mathbf{0} \Rightarrow \psi(y \equiv \Phi\theta_o) = A\Phi\theta_o = \left(\Phi^T\Phi\right)^{-1}\Phi^T\Phi\theta_o = \theta_o$
- In general, $e \neq \mathbf{0} \Rightarrow y = \Phi\theta_o + e \neq \Phi\theta_o \Rightarrow \hat{\theta} = Ay = \left(\Phi^T\Phi\right)^{-1}\Phi^T(\Phi\theta_o + e) = \theta_o + Ae \neq \theta_o$

Example: let $N =$ number of measurements $= 3$, $n =$ number of parameters $= 2$ and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N\colon \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon\right\}$ with known ε



- Each data vector y is an element of the **Measurements space** $\mathbb{R}^N = \mathbb{R}^3$
- If the least squares operator $A = \left(\Phi^T\Phi\right)^{-1}\Phi^T$ is applied as estimator ψ to the data vector y, then the estimate $\hat{\theta} = \psi\left(y\right) = Ay$ is an element of the **Parameter space** $\mathbb{R}^n = \mathbb{R}^2$ obtained as linear mapping of $y \in \mathbb{R}^N$ into $\mathbb{R}^n$
- If $e = \mathbf{0} \Rightarrow \psi\left(y \equiv \Phi\theta_o\right) = A\Phi\theta_o = \left(\Phi^T\Phi\right)^{-1}\Phi^T\Phi\theta_o = \theta_o$
- In general, $e \neq \mathbf{0} \Rightarrow y = \Phi\theta_o + e \neq \Phi\theta_o \Rightarrow \hat{\theta} = Ay = \left(\Phi^T\Phi\right)^{-1}\Phi^T(\Phi\theta_o + e) = \theta_o + Ae \neq \theta_o$
- Even if $\Phi\theta_o$ is unknown, its ℓ_∞–distance from y is unknown but bounded by ε: $y - \Phi\theta_o = e \in \mathcal{B}_e^\infty$

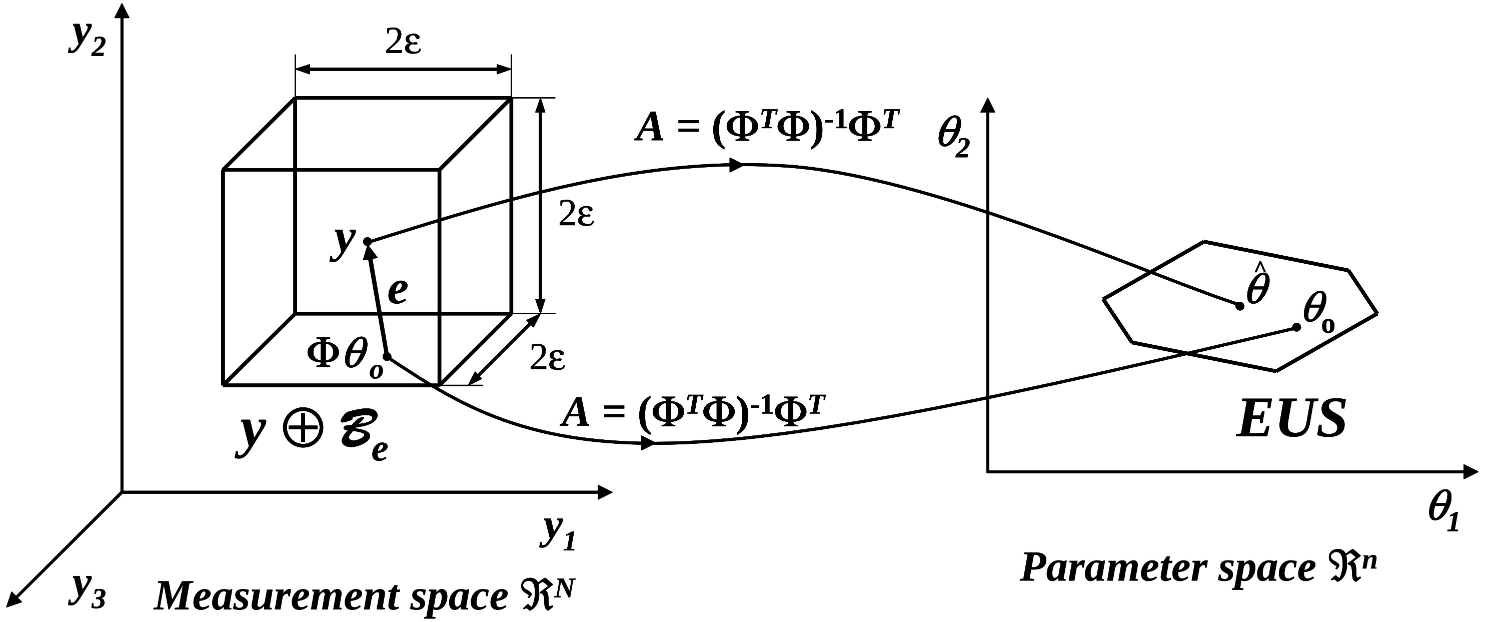
Example: let N = number of measurements $= 3$, n = number of parameters $= 2$ and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{ \tilde{e} \in \mathbb{R}^N \colon \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \leq \varepsilon \right\}$ with known ε



- Even if $\Phi\theta_o$ is unknown, its ℓ_∞–distance from y is unknown but bounded by ε: $y - \Phi\theta_o = e \in \mathcal{B}_e^\infty$ $\Rightarrow$ $\Phi\theta_o$ is an element of the **Measurement Uncertainty Set** MUS^∞ that contains any possible data vector whose ℓ_∞–distance from y is bounded by ε (MUS^∞ is convex and symmetric w.r.t. y):

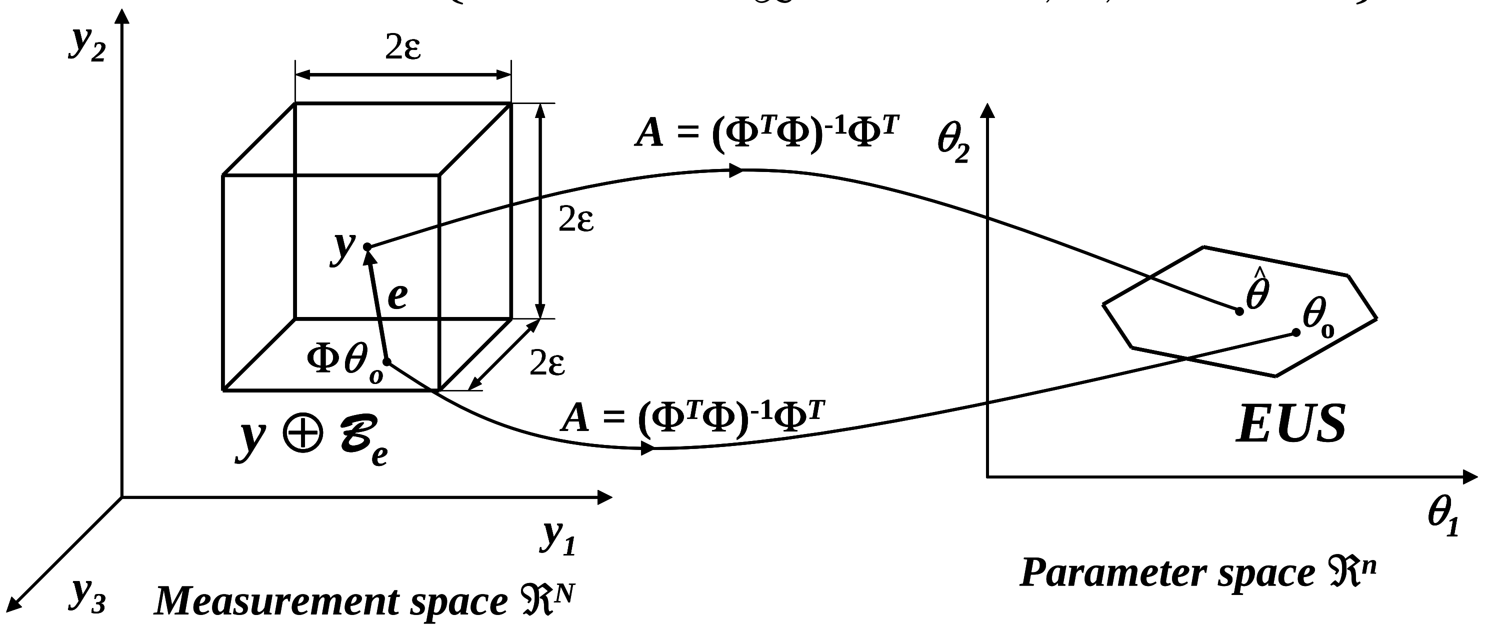
$$MUS^\infty = y \oplus \mathcal{B}_e^\infty = \{\tilde{y} \in \mathbb{R}^N : \|\tilde{y} - y\|_\infty = \max_{i=1,\ldots,N} |\tilde{y}_i - y_i| \leq \varepsilon\} \subset \mathbb{R}^N$$

Example: let N = number of measurements = 3, n = number of parameters = 2 and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N \colon \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \le \varepsilon\right\}$ with known ε



- Even if $\Phi\theta_o$ is unknown, its ℓ_∞–distance from y is unknown but bounded by ε: $y - \Phi\theta_o = e \in \mathcal{B}_e^\infty$ $\Rightarrow$ $\Phi\theta_o$ is an element of the **Measurement Uncertainty Set** MUS^∞ that contains any possible data vector whose ℓ_∞–distance from y is bounded by ε (MUS^∞ is convex and symmetric w.r.t. y):
$$MUS^\infty = y \oplus \mathcal{B}_e^\infty = \{\tilde{y} \in \mathbb{R}^N : \|\tilde{y} - y\|_\infty = \max_{i=1,\ldots,N} |\tilde{y}_i - y_i| \le \varepsilon\} \subset \mathbb{R}^N$$
- If the least squares operator A is applied as estimator ψ to all the elements of the MUS^∞, then the **Estimate Uncertainty Set** EUS^∞ is derived:
$$EUS^\infty = \psi\left(MUS^\infty\right) = A[MUS^\infty] = A[y \oplus \mathcal{B}_e^\infty] = Ay \oplus A[\mathcal{B}_e^\infty] = \hat{\theta} \oplus A[\mathcal{B}_e^\infty]$$

Example: let N = number of measurements $= 3$, n = number of parameters $= 2$ and assume that $y = \Phi\theta_o + e$, where $e \in \mathcal{B}_e^\infty = \left\{\tilde{e} \in \mathbb{R}^N : \|\tilde{e}\|_\infty = \max_{i=1,\ldots,N} |\tilde{e}_i| \le \varepsilon\right\}$ with known ε



- Even if $\Phi\theta_o$ is unknown, its ℓ_∞–distance from y is unknown but bounded by ε: $y - \Phi\theta_o = e \in \mathcal{B}_e^\infty$ $\Rightarrow$ $\Phi\theta_o$ is an element of the **Measurement Uncertainty Set** MUS^∞ that contains any possible data vector whose ℓ_∞–distance from y is bounded by ε (MUS^∞ is convex and symmetric w.r.t. y):
$$MUS^\infty = y \oplus \mathcal{B}_e^\infty = \{\tilde{y} \in \mathbb{R}^N : \|\tilde{y} - y\|_\infty = \max_{i=1,\ldots,N} |\tilde{y}_i - y_i| \le \varepsilon\} \subset \mathbb{R}^N$$
- If the least squares operator A is applied as estimator ψ to all the elements of the MUS^∞, then the **Estimate Uncertainty Set** EUS^∞ is derived:
$$EUS^\infty = \psi\left(MUS^\infty\right) = A[MUS^\infty] = A[y \oplus \mathcal{B}_e^\infty] = Ay \oplus A[\mathcal{B}_e^\infty] = \hat{\theta} \oplus A[\mathcal{B}_e^\infty]$$
- EUS^∞ is convex and symmetric with respect to $\hat{\theta}$; $\Phi\theta_o \in MUS^\infty \Rightarrow \theta_o = A\Phi\theta_o \in EUS^\infty$

- The EUS “volume” gives an idea of the estimation “quality” and, in particular, the **Estimate Uncertainty Intervals** $EUI_j, j=1,\ldots,n$, provide this measure:

$$EUI_j = \Big[\underbrace{\min_{\theta \in EUS} \theta_j}_{\hat{\theta}_j^m}, \underbrace{\max_{\theta \in EUS} \theta_j}_{\hat{\theta}_j^M}\Big] = \Big[\hat{\theta}_j^m, \hat{\theta}_j^M\Big] \subset \mathbb{R}$$

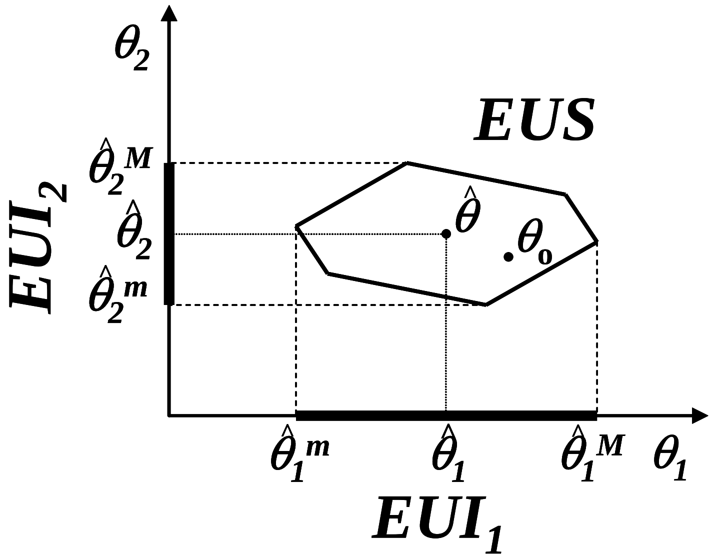
  – the range of the j-th component of the estimate is such that:

$$\hat{\theta}_j^m \le [\theta_o]_j \le \hat{\theta}_j^M$$

  – an upper bound on the estimation error of the j-th component is:

$$\left|\hat{\theta}_j - [\theta_o]_j\right| \le \left(\hat{\theta}_j^M - \hat{\theta}_j^m\right)\Big/ 2$$

  – $\hat{\theta}$ is the symmetry center of EUS, because EUS is the image of a symmetric set under a linear mapping



Set Membership framework: the linear estimation problem

- **“A priori” information**:
  - a deterministic source of data, that linearly depends on the unknown deterministic parameter $\theta_o \in \mathbb{R}^n$, generates the measurements:

$$y = \Phi\theta_o + e$$

    where the matrix $\Phi \in \mathbb{R}^{N \times n}$ is known
  - noise e is unknown but bounded: $e \in \mathcal{B}_e$, where the set $\mathcal{B}_e \subset \mathbb{R}^N$ is known
- **“A posteriori” information**:
  - a data vector $y \in \mathbb{R}^N$ is available
- **Goal:** find a suitable estimate $\hat{\theta} \in \mathbb{R}^n$ of the unknown deterministic parameter θ_o

$$\hat{\theta} = \psi(y) \simeq \theta_o$$

  and evaluate the estimate uncertainty in two cases: $e \in \mathcal{B}_e^\infty$ and $e \in \mathcal{B}_e^2$
- **Possible solution:** use the least squares estimator $\Rightarrow \hat{\theta} = Ay = \left(\Phi^T\Phi\right)^{-1}\Phi^T y$

Evaluation of $\mathbf{EUS}^\infty$

- The uncertainty set is a cube in $\mathbb{R}^N$ centered in the origin:

$$\mathcal{B}_e^\infty = \left\{ \tilde{e} \in \mathbb{R}^N : |\tilde{e}_i| \leq \varepsilon, i = 1, \ldots, N \right\}$$

$$\Downarrow \quad y = \Phi\theta_o + e$$

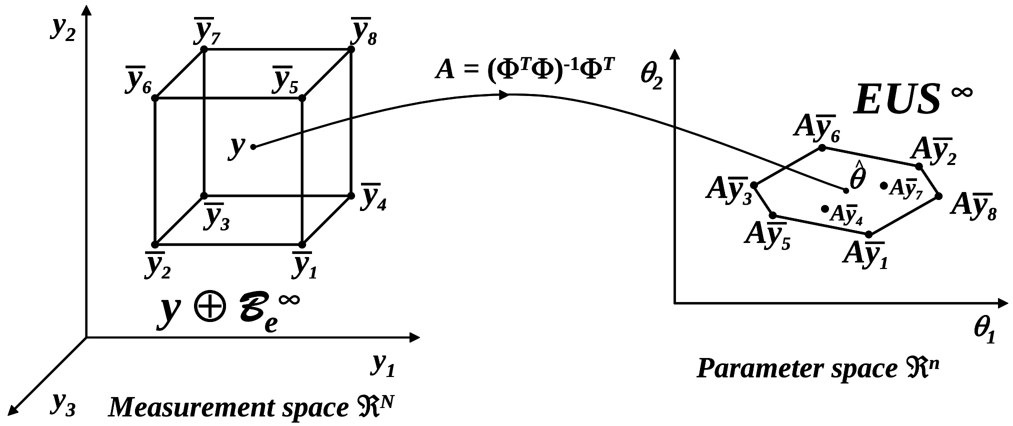
the set of any possible measurement (called **Measurement Uncertainty Set**) is a cube in $\mathbb{R}^N$ whose symmetry center is the data vector y:

$$MUS^\infty = y \oplus \mathcal{B}_e^\infty = \left\{ \tilde{y} \in \mathbb{R}^N : |\tilde{y}_i - y_i| \leq \varepsilon, i = 1, \ldots, N \right\} \subset \mathbb{R}^N$$

the vertices of MUS^∞ are denoted by $\bar{y}_k$, $k = 1, \ldots, 2^N$

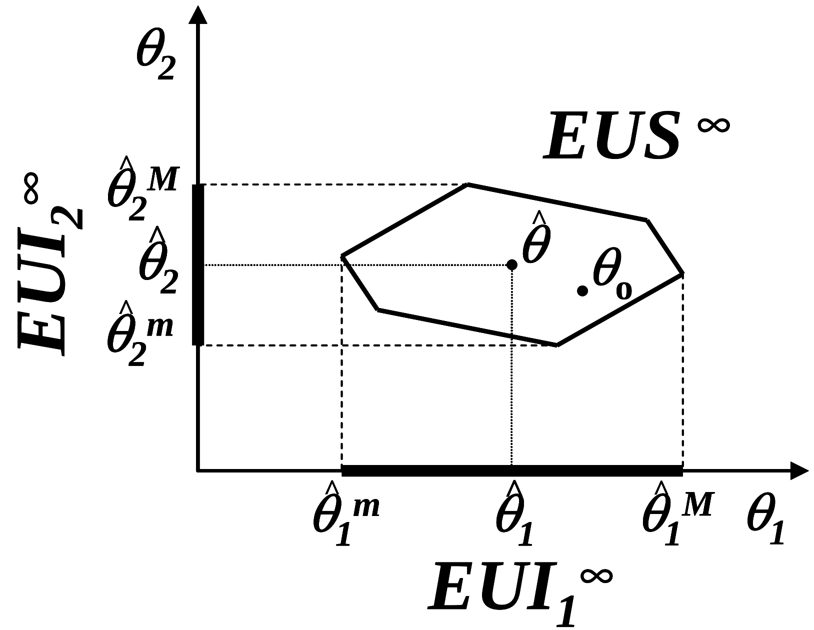
- **Theorem:** $EUS^\infty = A[MUS^\infty] = \text{conv}\left\{ A\bar{y}_k, k = 1, \ldots, 2^N \right\} \subset \mathbb{R}^n$

$\text{conv}\{\theta_1, \ldots, \theta_p\}$:
convex hull of the set $\{\theta_1, \ldots, \theta_p\}$
is the smallest convex polyhedron
(polytope) containing $\theta_1, \ldots, \theta_p$



Evaluation of $\mathbf{EUI}_j^\infty$

- **Theorem:** $EUI_j^\infty = \left[\hat{\theta}_j^m, \hat{\theta}_j^M\right] \subset \mathbb{R}$



where $\hat{\theta}_j^m = \sum_{k=1}^{N} a_{jk} \left[y_k - \varepsilon \cdot \operatorname{sign}\left(a_{jk}\right)\right]$,

$$\hat{\theta}_j^M = 2\hat{\theta}_j - \hat{\theta}_j^m,$$

$$A = \left[a_{jk}\right] = (\Phi^T \Phi)^{-1} \Phi^T,$$

$$\hat{\theta} = \left[\hat{\theta}_j\right] = Ay$$

Proof: $\hat{\theta}_j^m = \min_{\theta \in EUS^\infty} \theta_j = \min_{\tilde{y} \in MUS^\infty} (A\tilde{y})_j =$

$$= \min_{\substack{\tilde{y}: |\tilde{y}_i - y_i| \le \varepsilon \\ i=1,\ldots,N}} \sum_{k=1}^{N} a_{jk}\tilde{y}_k = \min_{\substack{\tilde{y}: -\varepsilon \le \tilde{y}_i - y_i \le \varepsilon \\ i=1,\ldots,N}} \sum_{k=1}^{N} a_{jk}\tilde{y}_k = \min_{\substack{\tilde{y}: y_i - \varepsilon \le \tilde{y}_i \le y_i + \varepsilon \\ i=1,\ldots,N}} \sum_{k=1}^{N} a_{jk}\tilde{y}_k$$

and such a minimum is achieved by $\tilde{y}_k = y_k - \varepsilon$ if $a_{jk} > 0$, or by $\tilde{y}_k = y_k + \varepsilon$ if $a_{jk} < 0$.

Since $MUS^\infty = y \oplus \mathcal{B}_e^\infty$ is symmetric with respect to the data vector y, then $EUS^\infty = A\left[MUS^\infty\right]$ is symmetric with respect to the estimate $\hat{\theta} = Ay$ and then:

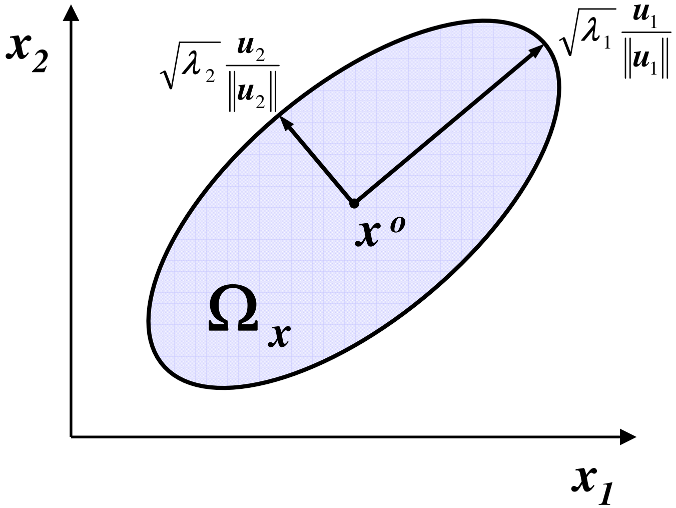
$$\hat{\theta}_j = \left(\hat{\theta}_j^m + \hat{\theta}_j^M\right) \Big/ 2, j = 1, \ldots, n \quad \Rightarrow \quad \hat{\theta}_j^M = 2\hat{\theta}_j - \hat{\theta}_j^m, j = 1, \ldots, n \quad \blacksquare$$

Description of ellipsoids

Let Ω_x be an ellipsoid in $\mathbb{R}^N$ centered in x^o:

$$\Omega_x = \left\{ x \in \mathbb{R}^N : (x - x^o)^T \Sigma_x^{-1} (x - x^o) \leq 1 \right\}$$

- The form matrix $\Sigma_x \in \mathbb{R}^{N \times N}$ is symmetric and positive definite $\Rightarrow$ it is invertible
- The directions of the main axes of Ω_x are given by the eigenvectors u_i of Σ_x, which are orthogonal because Σ_x is positive definite
- The lengths of the semi-axes of Ω_x are given by $\sqrt{\lambda_i(\Sigma_x)}$, where $\lambda_i(\Sigma_x)$ is the i-th eigenvalue of Σ_x



Linear transformation of ellipsoids

Let Ω_x be an ellipsoid in $\mathbb{R}^N$ centered in x^o:

$$\Omega_x = \left\{ x \in \mathbb{R}^N : (x - x^o)^T \Sigma_x^{-1} (x - x^o) \leq \varepsilon^2 \right\}$$

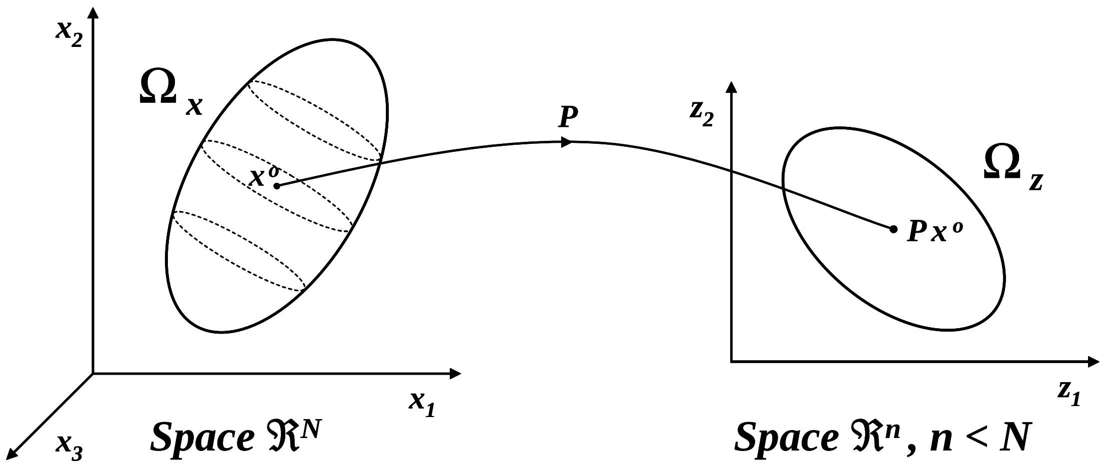
and consider the linear transformation:

$$z = Px \in \mathbb{R}^n, \text{ with } P \in \mathbb{R}^{n \times N}, n < N$$

- **Theorem:** if $\mathrm{rank}\,(P) = n$, then

$$\Omega_z = P\left[\Omega_x\right] = \left\{ z \in \mathbb{R}^n : (z - z^o)^T \Sigma_z^{-1} (z - z^o) \leq \varepsilon^2 \right\}$$

$$z^o = Px^o \in \mathbb{R}^n, \quad \Sigma_z = P\Sigma_x P^T \in \mathbb{R}^{n \times n}$$



Evaluation of $\mathbf{EUS}^2$

- The uncertainty set is a sphere in $\mathbb{R}^N$ centered in the origin:

$$\mathcal{B}_e^2 = \left\{ \tilde{e} \in \mathbb{R}^N : \tilde{e}^T \cdot \tilde{e} \leq \varepsilon^2 \right\}$$

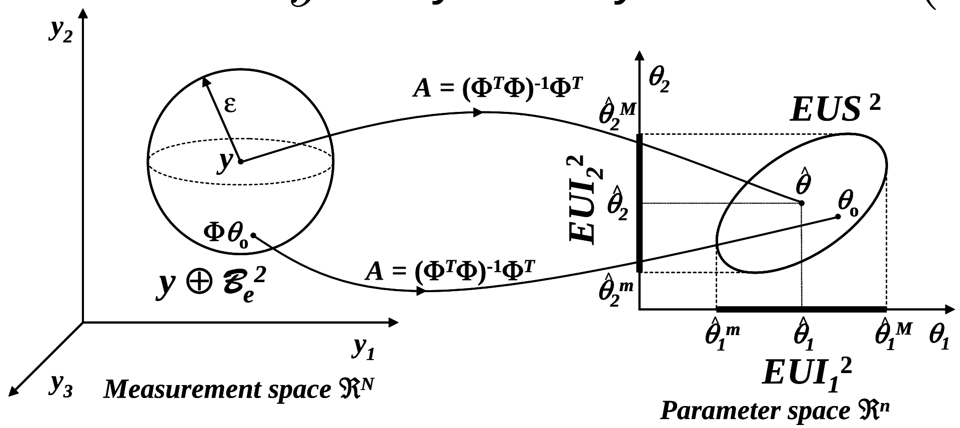
$$\Downarrow \qquad y = \Phi\theta_o + e$$

the set of any possible measurement (called **Measurement Uncertainty Set**) is a sphere in $\mathbb{R}^N$ whose symmetry center is the data vector y:

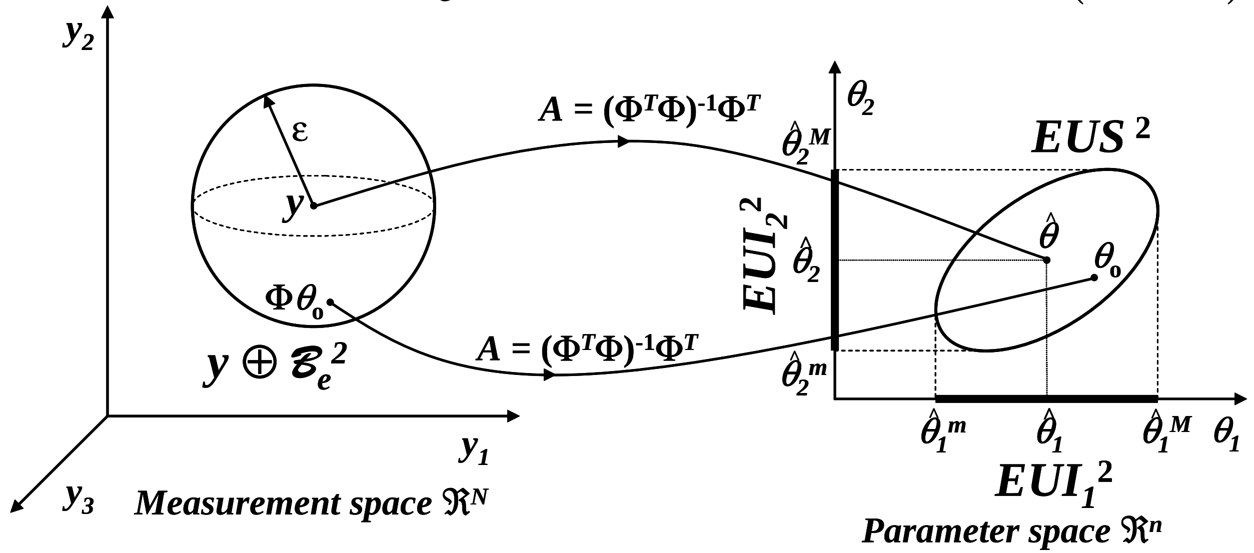
$$MUS^{\,2} = y \oplus \mathcal{B}_e^2 = \left\{ \tilde{y} \in \mathbb{R}^N : (\tilde{y} - y)^T \cdot (\tilde{y} - y) \leq \varepsilon^2 \right\} \subset \mathbb{R}^N$$

- **Theorem**: $EUS^2 = A\left[y \oplus \mathcal{B}_e^2\right] = \left\{ \tilde{\theta} \in \mathbb{R}^n : (\tilde{\theta} - \hat{\theta})^T \Phi^T \Phi \, (\tilde{\theta} - \hat{\theta}) \leq \varepsilon^2 \right\} \subset \mathbb{R}^n$

is an ellipsoid in $\mathbb{R}^n$ with $\hat{\theta} = Ay$ as symmetry center and $(\Phi^T\Phi)^{-1}$ as form matrix



- **Theorem**: $EUS^2 = A\left[y \oplus \mathcal{B}_e^2\right] = \left\{\tilde{\theta} \in \mathbb{R}^n : (\tilde{\theta} - \hat{\theta})^T \Phi^T \Phi\, (\tilde{\theta} - \hat{\theta}) \leq \varepsilon^2\right\} \subset \mathbb{R}^n$ is an ellipsoid in $\mathbb{R}^n$ with $\hat{\theta} = Ay$ as symmetry center and $(\Phi^T \Phi)^{-1}$ as form matrix



Proof: by definition, $EUS^{\,2}$ is the linear mapping of $MUS^{\,2} = y \oplus \mathcal{B}_e^2$ through the matrix A:

$$EUS^{\,2} = A\left[y \oplus \mathcal{B}_e^2\right] = \left\{\tilde{\theta} \in \mathbb{R}^n : (\tilde{\theta} - Ay)^T [AA^T]^{-1} (\tilde{\theta} - Ay) \leq \varepsilon^2\right\}$$

But $Ay = \hat{\theta}$, $A = (\Phi^T \Phi)^{-1} \Phi^T$ and then:

$$AA^T = (\Phi^T \Phi)^{-1} \Phi^T [(\Phi^T \Phi)^{-1} \Phi^T]^T = (\Phi^T \Phi)^{-1} \Phi^T \left\{\Phi [(\Phi^T \Phi)^{-1}]^T\right\} =$$

$$= \underbrace{(\Phi^T \Phi)^{-1} \Phi^T \Phi}_{I}\, [(\Phi^T \Phi)^T]^{-1} = (\Phi^T \Phi)^{-1}$$

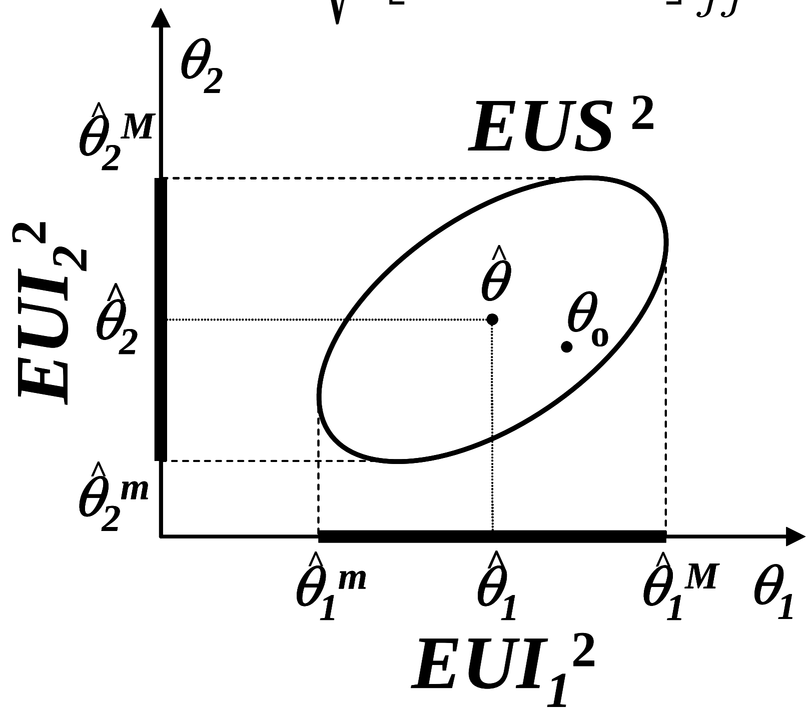
■

Evaluation of $\mathbf{EUI}_j^2$

- **Theorem**:

$$EUI_j^2 = \Big[\underbrace{\hat{\theta}_j - \varepsilon \cdot \sigma_j}_{\hat{\theta}_j^{\,m}}, \underbrace{\hat{\theta}_j + \varepsilon \cdot \sigma_j}_{\hat{\theta}_j^{\,M}}\Big] = \left[\hat{\theta}_j^{\,m}, \hat{\theta}_j^{\,M}\right] \subset \mathbb{R}$$

$$\sigma_j = \sqrt{\left[\left(\Phi^T\Phi\right)^{-1}\right]_{jj}}$$



Optimal (with minimal uncertainty) estimates

- Is the EUS the smallest set containing the “true” parameter θ_o?
- Are the EUI_j the smallest possible uncertainty intervals?
- Does the LS estimator provide the minimal uncertainty intervals?

To answer all these questions, it is necessary to analyze the set of all the parameters that are consistent with both the data and the available information on noise

- Definition: a parameter $\tilde{\theta}$ is said to be **feasible** (or consistent) if $(y - \Phi\tilde{\theta}) \in \mathcal{B}_e$

$$\Downarrow$$

$$FPS = \left\{\tilde{\theta} \in \mathbb{R}^n : (y - \Phi\tilde{\theta}) \in \mathcal{B}_e\right\} = \textbf{Feasible Parameter Set} =$$

$=$ set of all the parameters consistent with both the data and the information on noise and on the estimation problem

- FPS is independent of the estimation algorithm
- If data are generated by the “true” parameter θ_o, then θ_o is feasible; in fact:

$$y = \Phi\theta_o + e, e \in \mathcal{B}_e \quad \Rightarrow \quad y - \Phi\theta_o = e \in \mathcal{B}_e \quad \Rightarrow \quad \theta_o \in FPS$$

Relationship between $\mathbf{FPS}$ and $\mathbf{EUS}$

- **Theorem**:

$$\boxed{FPS \subseteq EUS}$$

Proof: if $\tilde{\theta} \in FPS$, then

$$(y - \Phi\tilde{\theta}) \in \mathcal{B}_e \quad \Rightarrow \quad \Phi\tilde{\theta} \in y \oplus \mathcal{B}_e \quad \Rightarrow \quad A\left[\Phi\tilde{\theta}\right] \in A\left[y \oplus \mathcal{B}_e\right] = EUS$$

But $A\left[\Phi\tilde{\theta}\right] = (\Phi^T\Phi)^{-1}\Phi^T\,\Phi\tilde{\theta} = \tilde{\theta}$ and then $\tilde{\theta} \in EUS$. ■

- The **Parameter Uncertainty Intervals** PUI_j, $j = 1, \ldots, n$ are defined as:

$$PUI_j = \Big[\underbrace{\min_{\theta \in FPS} \theta_j}_{\theta_j^m}, \underbrace{\max_{\theta \in FPS} \theta_j}_{\theta_j^M}\Big] = \left[\theta_j^m, \theta_j^M\right] \subset \mathbb{R}$$

from the above theorem:

$$PUI_j \subseteq EUI_j,\ j = 1, \ldots, n$$

$$\Downarrow$$

$$\hat{\theta}_j^m \leq \theta_j^m \leq [\theta_o]_j \leq \theta_j^M \leq \hat{\theta}_j^M$$

Evaluation of $\mathbf{FPS}^{\infty}$ and $\mathbf{PUI}_j^{\infty}$

- If $\tilde{\theta} \in FPS^{\infty}$, then $(y - \Phi\tilde{\theta}) \in \mathcal{B}_e^{\infty} = \left\{ \tilde{e} \in \mathbb{R}^N : |\tilde{e}_i| \leq \varepsilon, i = 1, \ldots, N \right\}$

$$\Downarrow \quad \varphi_i^T : \text{ } i\text{-th row of } \Phi$$

$$\left|(y - \Phi\tilde{\theta})_i\right| = \left|y_i - \varphi_i^T \tilde{\theta}\right| \leq \varepsilon, \quad i = 1, \ldots, N \qquad \Rightarrow$$

$$FPS^{\infty} = \left\{ \tilde{\theta} \in \mathbb{R}^n : \left|y_i - \varphi_i^T \tilde{\theta}\right| \leq \varepsilon, i = 1, \ldots, N \right\}$$

i.e., FPS^{∞} is a polytope (a convex polyhedron) generated by linear inequalities:

$$\left|y_i - \varphi_i^T \tilde{\theta}\right| \leq \varepsilon \iff -\varepsilon \leq y_i - \varphi_i^T \tilde{\theta} \leq \varepsilon \iff y_i - \varepsilon \leq \varphi_i^T \tilde{\theta} \leq y_i + \varepsilon$$

- Moreover, $PUI_j^{\infty} = \Big[\underbrace{\min_{\theta \in FPS^{\infty}} \theta_j}_{\theta_j^m}, \underbrace{\max_{\theta \in FPS^{\infty}} \theta_j}_{\theta_j^M} \Big] = \left[\theta_j^m, \theta_j^M \right] \subset \mathbb{R}$

with θ_j^m and θ_j^M solutions of linear programming problems of the standard form:

$$\min_x c^T x \quad \text{with the constraint: } Ax \leq b$$

Evaluation of $\mathbf{FPS}^2$ and $\mathbf{PUI}_j^2$

- **Theorem:** $FPS^2 = \left\{ \tilde{\theta} \in \mathbb{R}^n : (\tilde{\theta} - \hat{\theta})^T \left[\Phi^T \Phi\right] (\tilde{\theta} - \hat{\theta}) \leq \varepsilon^2 - \alpha^2 \right\}$

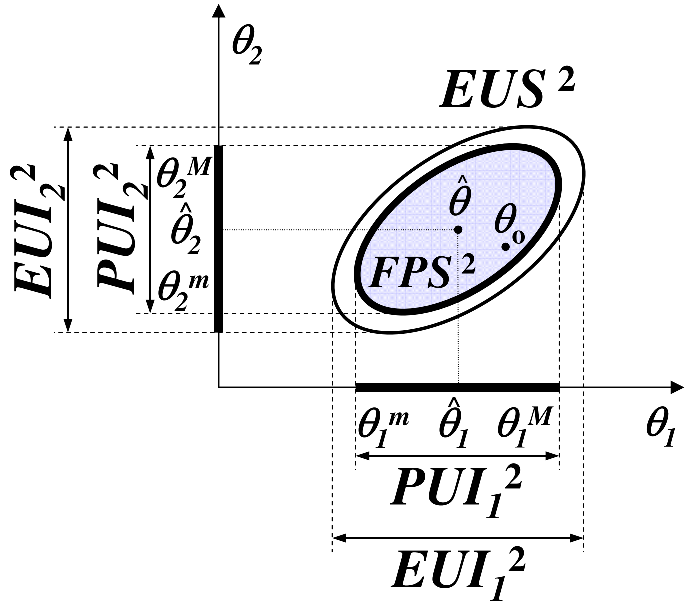
$$\alpha^2 = (y - \Phi\hat{\theta})^T (y - \Phi\hat{\theta}) = \| y - \Phi\hat{\theta} \|_2^2 \leq \varepsilon^2$$

$=$ "fitting error" between measured outputs and estimated outputs

a greater fitting error $\Rightarrow$ a smaller FPS^2 $\Rightarrow$ a lower uncertainty on parameters

- Moreover, $PUI_j^2 = \Big[\underbrace{\hat{\theta}_j - \sigma_j \sqrt{\varepsilon^2 - \alpha^2}}_{\theta_j^m}, \underbrace{\hat{\theta}_j + \sigma_j \sqrt{\varepsilon^2 - \alpha^2}}_{\theta_j^M} \Big] = \left[\theta_j^m, \theta_j^M\right] \subset \mathbb{R}$

$$\sigma_j = \sqrt{\left[(\Phi^T \Phi)^{-1}\right]_{jj}}$$



Optimal estimates

- Definition: given an estimate $\hat{\theta}$, the **estimate error** $\mathcal{E}(\hat{\theta})$ is given by:

$$\mathcal{E}(\hat{\theta}) = \sup_{\theta \in FPS} \| \theta - \hat{\theta} \|$$

- Definition: an estimate $\hat{\theta}^{opt}$ is **optimal** if:

$$\mathcal{E}(\hat{\theta}^{opt}) \leq \mathcal{E}(\hat{\theta}), \quad \forall \hat{\theta} \in \mathbb{R}^n$$

- **Central estimate**:

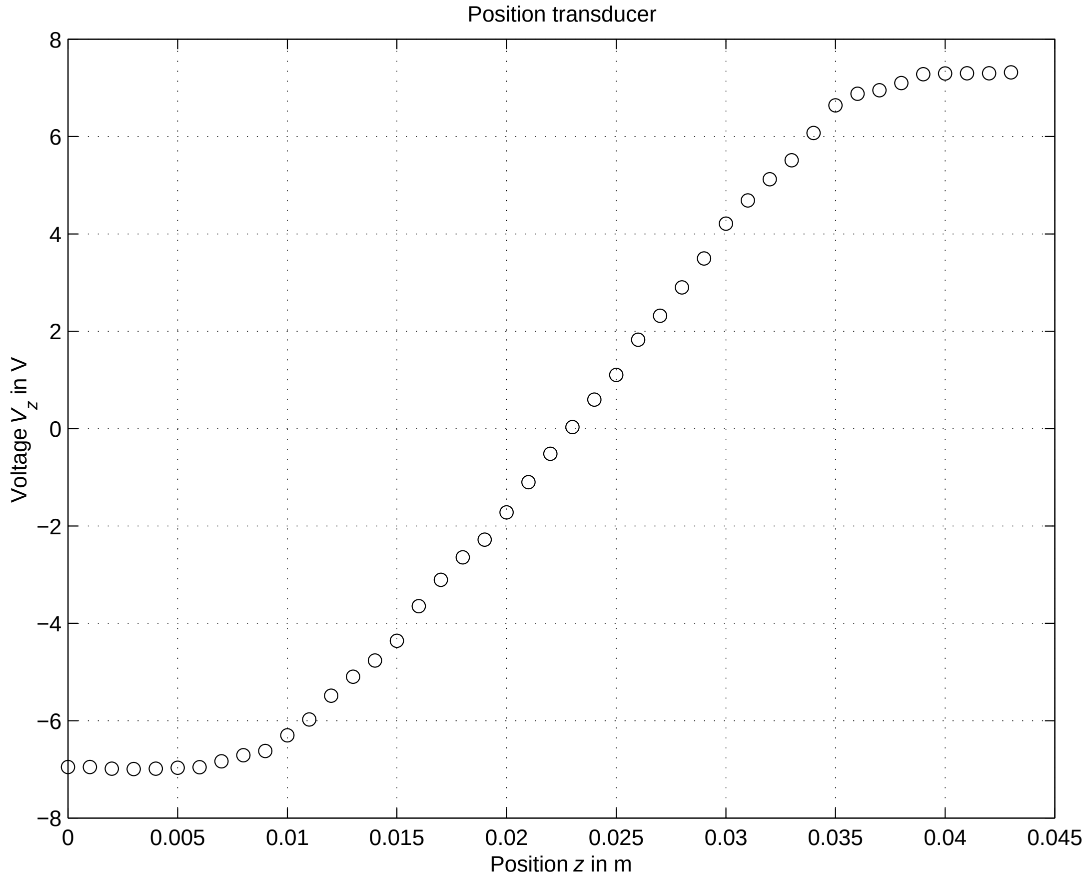
$$\hat{\theta}^C = \left[\hat{\theta}_j^C \right], \quad \text{where } \hat{\theta}_j^C = \left(\theta_j^m + \theta_j^M\right) \Big/ 2, \quad j = 1, \ldots, n$$

  - the central estimate is optimal both if $\mathcal{B}_e = \mathcal{B}_e^2$ and if $\mathcal{B}_e = \mathcal{B}_e^\infty$, since:

$$\left| [\theta_o]_j - \hat{\theta}_j^C \right| \leq \left(\theta_j^M - \theta_j^m\right) \Big/ 2, \quad j = 1, \ldots, n$$

  - if $\mathcal{B}_e = \mathcal{B}_e^2$, the least squares estimate $\hat{\theta}^{LS} = (\Phi^T \Phi)^{-1} \Phi^T y$ is central $\Rightarrow$ $\hat{\theta}^{LS}$ is optimal if $\mathcal{B}_e = \mathcal{B}_e^2$, but in general it is not optimal if $\mathcal{B}_e = \mathcal{B}_e^\infty$

Example: parametric estimation of a position transducer model



The static characteristic of the position-voltage transducer is nearly linear in the range between 1.3 e 3.5 cm $\Rightarrow$ the characteristic can be linearly approximated by:

$$V_z = K_t \cdot z + V_o$$

- In the linearity interval between 1.3 e 3.5 cm:

$$V_z = \underbrace{K_t}_{\text{unknown}} \cdot z + \underbrace{V_o}_{\text{unknown}}$$

- The most relevant error occurs in the position z measurement and it is not greater than 0.5 mm $\Rightarrow$ to account for this error, the model equation can be rewritten as:

$$z = \frac{1}{K_t} \cdot V_z - \frac{V_o}{K_t} + e$$

where the unknown parameters are:

$$\theta_1 = \frac{1}{K_t}, \quad \theta_2 = -\frac{V_o}{K_t}$$

- The N measurements taken in the linearity interval form a system of equations:

$$\begin{aligned} z_1 &= V_{z,1} \cdot \theta_1 + \theta_2 + e_1 \\ z_2 &= V_{z,2} \cdot \theta_1 + \theta_2 + e_2 \\ &\vdots \\ z_N &= V_{z,N} \cdot \theta_1 + \theta_2 + e_N \end{aligned}$$

$V_{z,i}$: voltage provided by the transducer when the position value is z_i

- In matrix form:

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} = \begin{bmatrix} V_{z,1} & 1 \\ V_{z,2} & 1 \\ \vdots & \vdots \\ V_{z,N} & 1 \end{bmatrix} \cdot \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

i.e., the estimation problem is in the standard form:

$$y = \Phi \cdot \theta + e$$

where $y \in \mathbb{R}^N$, $\Phi \in \mathbb{R}^{N\times 2}$, $e \in \mathbb{R}^N$ and the unknown is $\theta \in \mathbb{R}^2$

- Using the Least Squares estimation algorithm:

$$\hat{\theta} = A \cdot y, \quad \text{with } A = (\Phi^T \cdot \Phi)^{-1}\Phi^T \quad \Rightarrow$$

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \begin{bmatrix} 1.8194 \cdot 10^{-3} \\ 2.2791 \cdot 10^{-2} \end{bmatrix} \quad \Rightarrow$$

$$\hat{K}_t = \frac{1}{\hat{\theta}_1} = 549.62\,\mathrm{V/m}, \quad \hat{V}_o = -\frac{\hat{\theta}_2}{\hat{\theta}_1} = -12.526\,\mathrm{V}$$

Evaluation of the Estimate Uncertainty Intervals $\mathbf{EUI}_j^\infty$

$$e^N \in \mathcal{B}_e^\infty = \left\{\tilde{e}^N \in \mathbb{R}^N : |\tilde{e}_i| \leq \varepsilon,\ i = 1, \ldots, N\right\}, \quad \varepsilon = 5 \cdot 10^{-4} \quad \Rightarrow$$

$$EUI_j^\infty = \left[\hat{\theta}_j^m = \min_{\theta \in EUS^\infty} \theta_j, \hat{\theta}_j^M = \max_{\theta \in EUS^\infty} \theta_j\right], \quad j = 1, 2$$

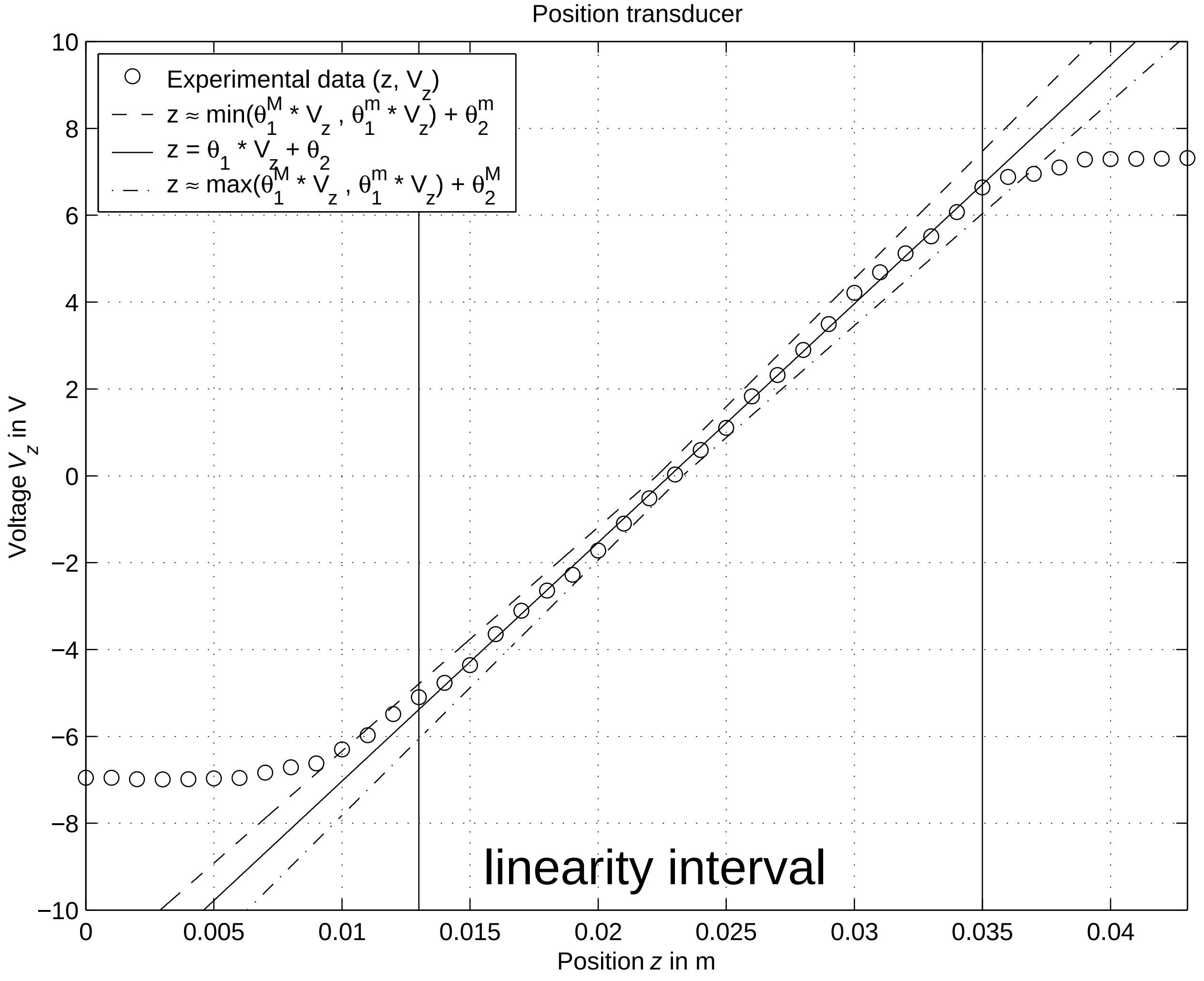
$$\begin{cases} \hat{\theta}_j^m = \min_{\theta \in EUS^\infty} \theta_j = \sum_{k=1}^N a_{jk} \cdot [y_k - \varepsilon \cdot \text{sign}(a_{jk})] \\ \hat{\theta}_j^M = \max_{\theta \in EUS^\infty} \theta_j = \sum_{k=1}^N a_{jk} \cdot [y_k + \varepsilon \cdot \text{sign}(a_{jk})] = 2 \cdot \hat{\theta}_j - \hat{\theta}_j^m \end{cases} \Rightarrow$$

$$\begin{aligned} \left[\hat{\theta}_1^m, \hat{\theta}_1^M\right] &= [1.6996 \cdot 10^{-3}, 1.9392 \cdot 10^{-3}] \\ \left[\hat{\theta}_2^m, \hat{\theta}_2^M\right] &= [2.2291 \cdot 10^{-2}, 2.3291 \cdot 10^{-2}] \end{aligned} \Rightarrow$$

$$\left[\hat{K}_t^m, \hat{K}_t^M\right] = \left[1/\hat{\theta}_1^M, 1/\hat{\theta}_1^m\right] = [515.67, 588.36] \text{ V/m}$$

$$\left[\hat{V}_o^m, \hat{V}_o^M\right] = \left[-\hat{\theta}_2^M/\hat{\theta}_1^m, -\hat{\theta}_2^m/\hat{\theta}_1^M\right] = [-13.703, -11.495] \text{ V}$$

Envelope of the static characteristics of models whose parameters θ are taken as the extremes of the Estimate Uncertainty Intervals $EUI_j^\infty, j = 1, 2$



Evaluation of the Parameter Uncertainty Intervals $\mathbf{PUI}_j^\infty$

$$FPS^\infty = \left\{ \tilde{\theta} \in \mathbb{R}^{dim(\tilde{\theta})} : \left| y_i - [\Phi \cdot \tilde{\theta}]_i \right| \leq \varepsilon,\ i = 1, \ldots, N \right\}$$

$$PUI_j^\infty = \left[\min_{\theta \in FPS^\infty} \theta_j, \max_{\theta \in FPS^\infty} \theta_j \right] \subseteq EUI_j^\infty, \quad j = 1, 2$$

The extremes of PUI_j^∞, $j = 1, 2$, are solutions of the linear programming problems

$$\begin{cases} \displaystyle\min_{\theta \in FPS^\infty} \theta_j = \min_{M \cdot \theta \leq b} c^T \theta \\ \displaystyle\max_{\theta \in FPS^\infty} \theta_j = -\min_{M \cdot \theta \leq b} (-c)^T \theta \end{cases} \qquad M = \begin{bmatrix} \Phi \\ -\Phi \end{bmatrix},\ b = \begin{bmatrix} y \\ -y \end{bmatrix} + \varepsilon,\ c = j\text{-th column of } I_{2\times 2}$$

$$\Downarrow$$

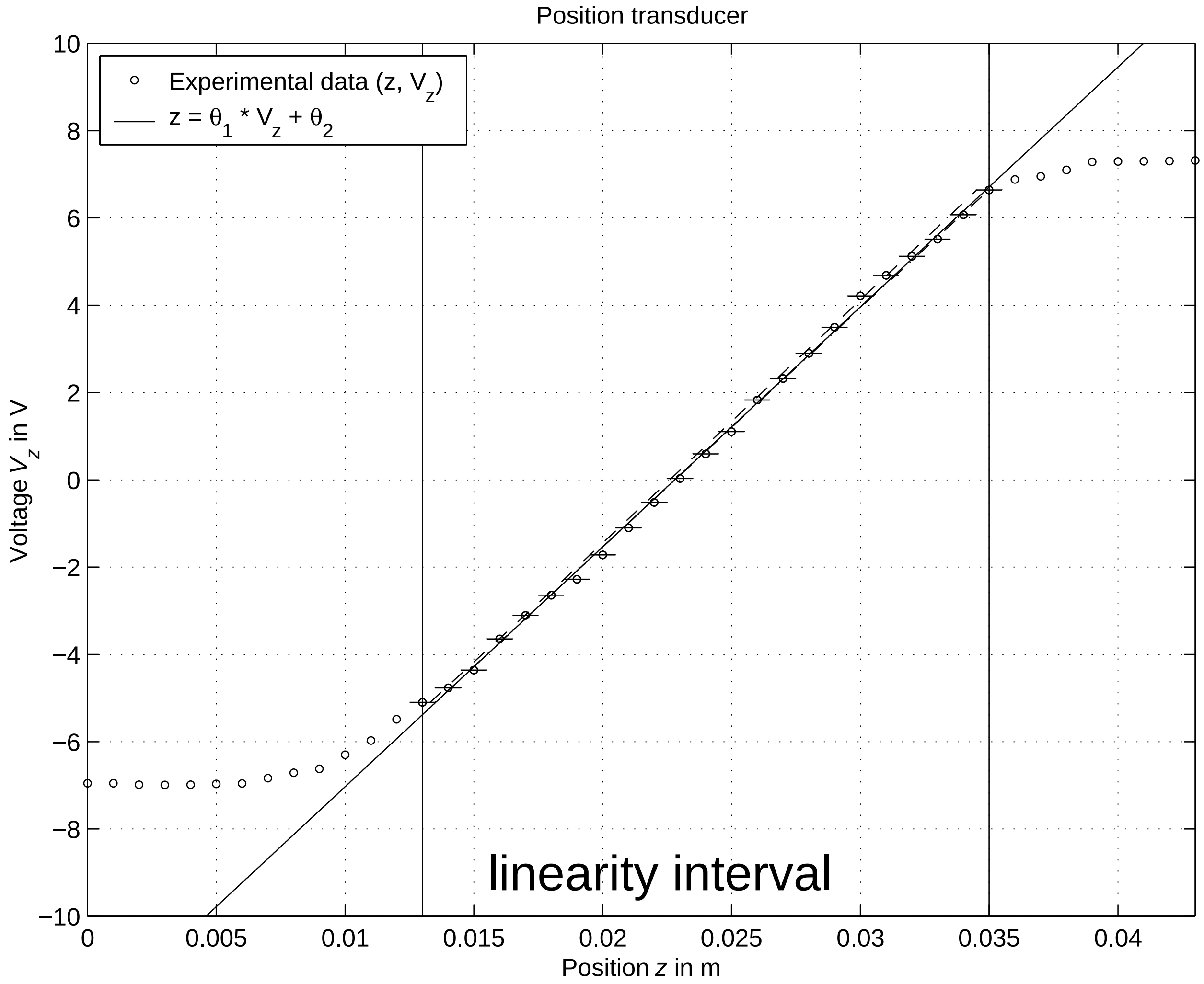
$$\left[\theta_1^m = \min_{\theta \in FPS^\infty} \theta_1, \theta_1^M = \max_{\theta \in FPS^\infty} \theta_1 \right] = [1.7909 \cdot 10^{-3}, 1.8484 \cdot 10^{-3}]$$

$$\left[\theta_2^m = \min_{\theta \in FPS^\infty} \theta_2, \theta_2^M = \max_{\theta \in FPS^\infty} \theta_2 \right] = [2.2596 \cdot 10^{-2}, 2.2807 \cdot 10^{-2}] \quad \Rightarrow$$

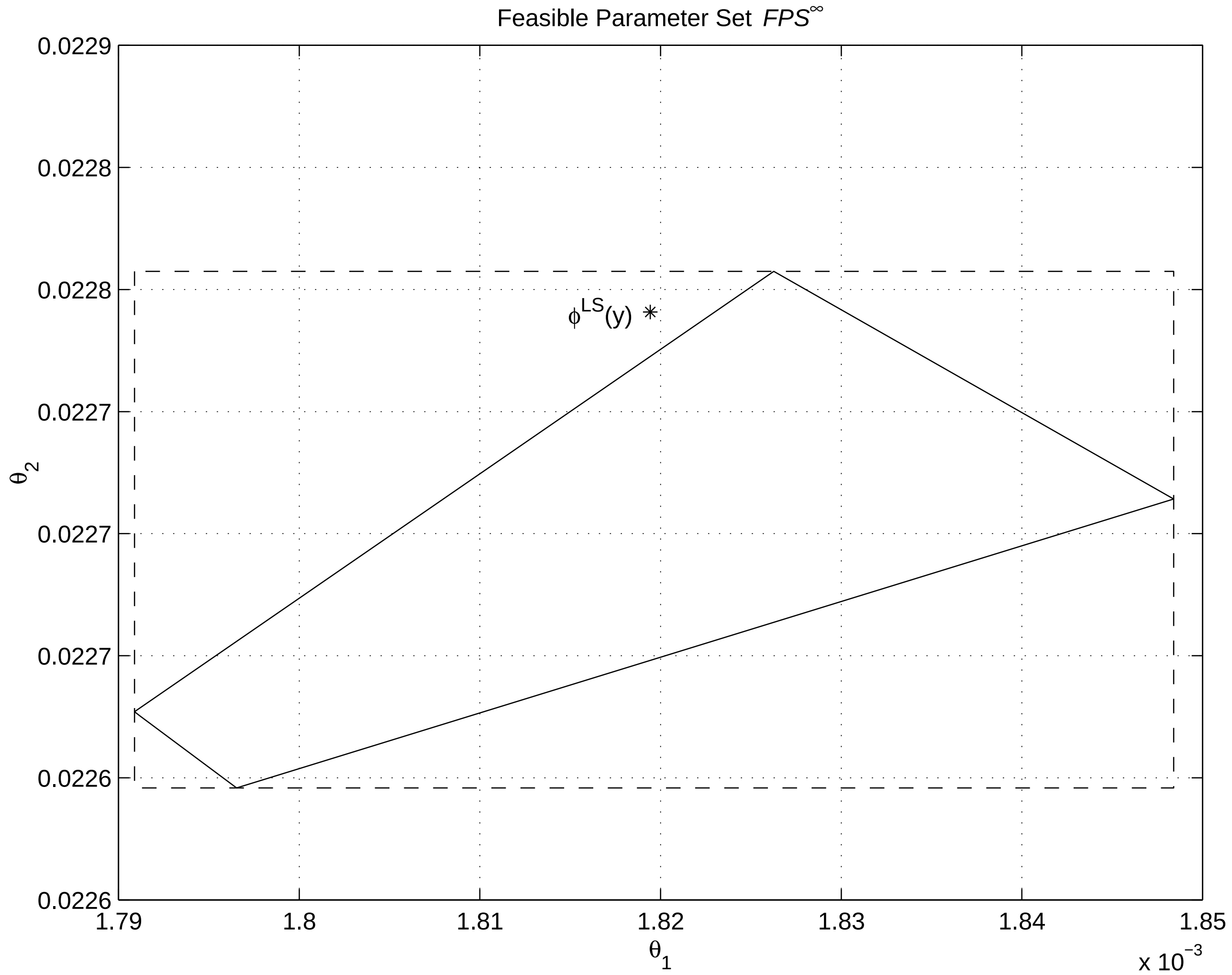
$$[K_t^m, K_t^M] = \left[1/\theta_1^M, 1/\theta_1^m \right] = [541.01, 558.38] \text{ V/m}$$

$$[V_o^m, V_o^M] = \left[-\theta_2^M/\theta_1^m, -\theta_2^m/\theta_1^M \right] = [-12.735, -12.225] \text{ V}$$

Envelope of the static characteristics of models whose parameters θ belong to the Feasible Parameter Set FPS^{∞}



Feasible Parameter Set FPS^{∞} (continuous line) and set of estimates given by the extremes of Parameter Uncertainty Intervals $PUI_j^{\infty}, j = 1, 2$



Essential references

F. C. Schweppe, *Uncertain Dynamics Systems*. Englewood Cliffs, NJ: Prentice Hall, 1973.

M. Milanese, R. Tempo, A. Vicino (editors), *Robustness in Identification and Control*. New York: Plenum Press, 1989.

M. Milanese, A. Vicino, “Optimal estimation theory for dynamic systems with set membership uncertainty: an overview,” *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.

Special Issue on System Identification for Robust Control Design, *IEEE Transactions on Automatic Control*, vol. AC-37, no. 7, pp. 899–974, 1992.

R. S. Smith, M. Dahleh (editors), *The Modeling of Uncertainty in Control Systems*, vol. 192 of *Lecture Notes in Control and Information Sciences*. London, UK: Springer-Verlag, 1994.

M. Milanese, J. Norton, H. Piet-Lahanier, É. Walter (editors), *Bounding Approaches to System Identification*. New York: Plenum Press, 1996.

J. R. Partington, *Interpolation, Identification, and Sampling*, vol. 17 of *London Mathematical Society Monographs New Series*. New York: Clarendon Press - Oxford, 1997.

A. Garulli, A. Tesi, A. Vicino (editors), *Robustness in Identification and Control*, vol. 245 of *Lecture Notes in Control and Information Sciences*. Godalming, UK: Springer-Verlag, 1999.

J. Chen, G. Gu, *Control-Oriented System Identification: An H_∞ Approach*. New York: John Wiley & Sons, Inc., 2000.