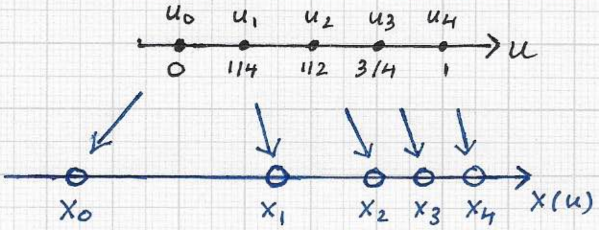
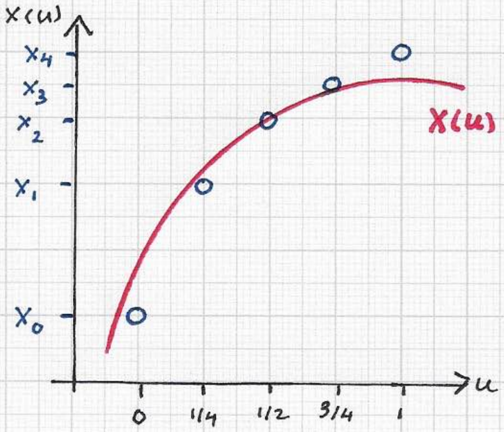
■ GEOMETRIC MODELING AND DENSITY – Cont'd.

• *Ex.:* In a 1D setting, one has to compute a best quadratic polynomial that, when evaluated for equidistantly spaced parameter values, generates function values that optimally approximate a given finite, discrete set of (feature) values. For example, given five feature values $x_0, \ldots, x_4$, associate five **equidistantly** spaced parameter values $u_0, u_1, \ldots, u_4$ with them and consider the conditions $x(u_i) = x_i$ for the computation of $x(u) = \sum_{j=0}^{2} c_j u^j$. The best quadratic polynomial function should produce x-value distributions that are "optimally similar" to the sample set $\{x_j\}_{j=0}^{4}$.





Best quadratic approximation

➔ The unique best quadratic approximation $x(u)$ defines a desired "model" for the given feature values. It is defined by the system $\sum_{j=0}^{2} c_j (u_i)^j = x_i$, $i = 0 \ldots 4$, or $c_0 + c_1 u_i + c_2 u_i^2 = x_i$.

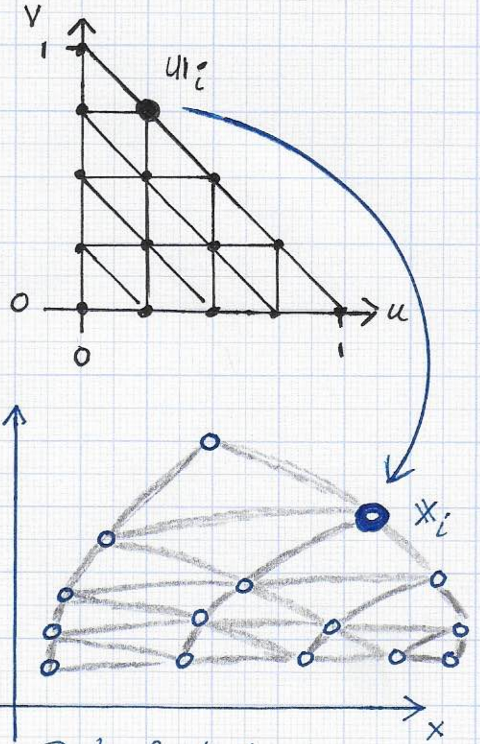
The overdetermined linear system is solved via the least-squares method, with minimal error $E_{RMS} = \left(\frac{1}{5} \sum_{i=0}^{4} (x_i - x(u_i))^2 \right)^{1/2}$.

$$\begin{bmatrix} 1 & u_0 & u_0^2 \\ & \cdots & \\ 1 & u_4 & u_4^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} x_0 \\ \cdots \\ x_4 \end{bmatrix}$$

$$\Longleftrightarrow \; U \mathbf{c} = \mathbf{x} \;\Rightarrow\; \mathbf{c} = (U^T U)^{-1} U^T \mathbf{x}.$$

GEOMETRIC MODELING AND DENSITY – Cont'd.

• **Ex.:** Considering the 2D case, the input to the best quadratic polynomial approximation step is a randomly distributed (feature) point set, $\{(x_i, y_i)^T\}_{i=0}^n$. (**Note**: Density values ρ_i are not given!) The goal is to map point-associated parameter values (tuples) (u_i, v_i) to their corresponding points in feature space. For example, one can use 15 equidistantly spaced tuples $(u_{i,j}, v_{i,j})$, $i,j \geq 0$ and $i+j \leq 4$, discretizing the "standard triangle" (left). Assuming that at least 15 feature points are given in xy-space, one must define a correspondence between tuples $\mathbb{U}_{i,j} = (u_{i,j}, v_{i,j})$ and points $\mathbb{X}_{i,j} = \ldots$. For simplicity, use a single-index notation, i.e., map a tuple $\mathbb{U}_i = (u_i, v_i)$ to point $\mathbb{X}_i = (x_i, y_i)^T$. (The 15 selected "correspondences" $\mathbb{U}_i \Longleftrightarrow \mathbb{X}_i$ suffice to compute the quadratic model.)



Data for best quadratic approximation

→ The best quadratic approximation (of the "space deformation") is determined by the conditions $\mathbb{X}(\mathbb{U}_i) = \mathbb{X}_i$ $(i=1 \ldots 15)$:

$$\mathbb{X}(\mathbb{U}_i) = \mathbb{X}(u_i, v_i) = (x(u_i,v_i), y(u_i,v_i))^T = \sum_{\substack{j,k \geq 0 \\ \wedge\, j+k \leq 2}} \mathbb{c}_{j,k}\, u_i^j v_i^k = \sum_{\|\mathbf{j}\| \leq 2} \mathbb{c}_{\mathbf{j}}\, \mathbb{U}_i^{\mathbf{j}} = \mathbb{X}_i$$

$$\Longleftrightarrow \mathbb{c}_{0,0} + \mathbb{c}_{1,0} u_i + \mathbb{c}_{0,1} v_i + \mathbb{c}_{1,1} u_i v_i + \mathbb{c}_{2,0} u_i^2 + \mathbb{c}_{0,2} v_i^2 = \mathbb{X}_i \quad \text{, where } \mathbb{c}_{j,k} = \begin{pmatrix} c^x_{j,k} \\ c^y_{j,k} \end{pmatrix}.$$

GEOMETRIC MODELING AND DENSITY – Cont'd.

• ... 2D example... The overdetermined linear system is:

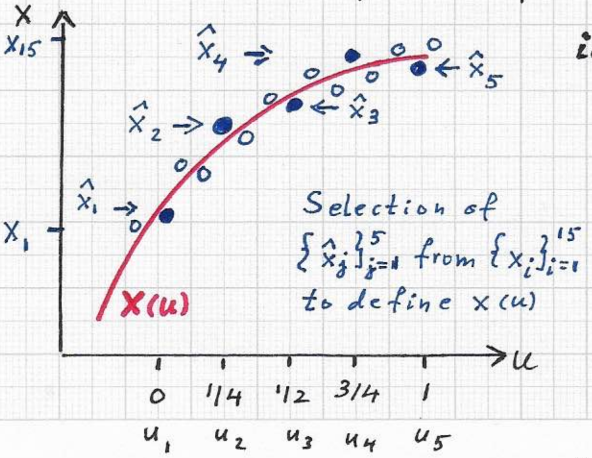
$$\begin{bmatrix} 1 & u_1 & v_1 & u_1 v_1 & u_1^2 & v_1^2 \\ \vdots & & & & & \vdots \\ 1 & u_{15} & v_{15} & u_{15} v_{15} & u_{15}^2 & v_{15}^2 \end{bmatrix} \begin{bmatrix} c_{0,0}^x & c_{0,0}^y \\ \vdots & \vdots \\ c_{0,2}^x & c_{0,2}^y \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_{15} & y_{15} \end{bmatrix}$$

$$\Leftrightarrow \quad U \quad C = X$$

$$\Rightarrow \quad C = (U^T U)^{-1} U^T X .$$

• **Note:** i) Determining an 'appropriate' correspondence between $\mathbf{u}_i$- and $\mathbf{x}_i$-values ("parametrization") is important for the error associated with $\mathbf{x}(\mathbf{u})$.

ii) Considering ALL given feature points $\mathbf{x}_i$, the best approximation (model) $\mathbf{x}(\mathbf{u})$ might not be good enough to generate feature point sets of very similar point distributions. In this case, splitting / subdivision of the feature point set can be used.



→ Alternative: choose averages of local x-value clusters to define $\hat{x}_j$.

iii) Consider the illustration (left). In this univariate example three coefficients must be determined for the quadratic polynomial $x(u)$. The number of equidistantly spaced parameter values is 5 ($5 \geq 3$); the number of (parametrized) given x-values is 15, and one can choose 5 of them to establish the overdetermined system for the coefficients.

Bernd Hamann — -14- — 5/25/2021

Stratovan

GEOMETRIC MODELING AND DENSITY – Cont'd.

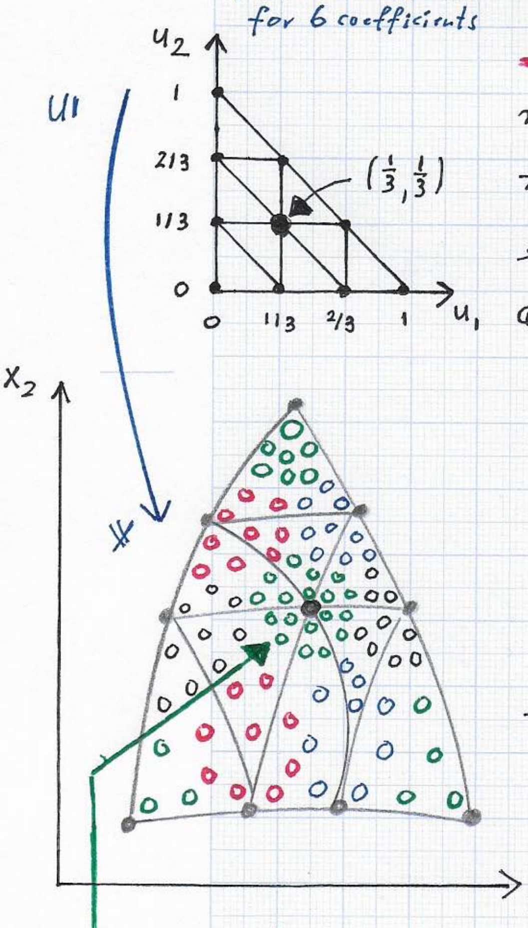
• **Note:** (cont'd.)

iv) One must consider the numbers of data involved in the computation of an approximation's coefficients. The general best quadratic approximation $\mathbb{X}(\mathbb{U})$,

$$\mathbb{X}(\mathbb{U}) = \sum_{\|\mathbb{j}\|\le 2}{}' \mathbb{c}_{\mathbb{j}}\, \mathbb{U}^{\mathbb{j}} = \sum_{\substack{j_1,\dots,j_D\ge 0 \\ \wedge\, j_1+\dots+j_D\le 2}}{}' \mathbb{c}_{j_1,\dots,j_D}\, u_1^{j_1}\cdots u_D^{j_D} ,$$

mapping a D-dimensional parameter tuple $(u_1,\dots,u_D)$ to a point $(x_1,\dots,x_D)^T$ in D-dimensional feature space, has $\frac{1}{2}(D+1)(D+2)$ coefficients $\mathbb{c}_{\mathbb{j}}$, $\mathbb{c}_{\mathbb{j}} = (c_{\mathbb{j}}^{x_1},\dots,c_{\mathbb{j}}^{x_D})^T$. Thus, at least $\frac{1}{2}(D+1)(D+2)$ conditions are required to define an overdetermined system for $\mathbb{X}(\mathbb{U})$'s coefficients. Assuming that the number of given (feature) points $\mathbb{X}_i$ is much larger than $\frac{1}{2}(D+1)(D+2)$, one can select an 'appropriate' subset of $\{\mathbb{X}_i\}$ to establish the linear system. (One can also use local means/averages of $\mathbb{X}_i$-clusters to set up the system.)

Bivariate example: set up 10 conditions for 6 coefficients



an $\mathbb{X}$-value selected from green cluster of points (or derived from this cluster) = image of parameter tuple $(\frac{1}{3},\frac{1}{3})$

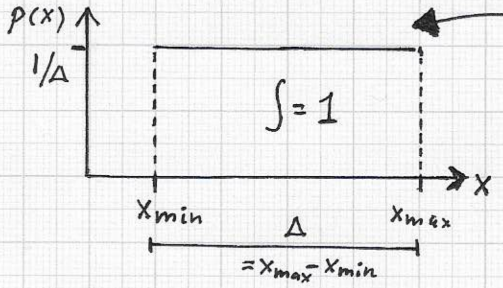
v) A best quadratic approximation $\mathbb{X}(\mathbb{U})$ is a 'good model' of the given data $\{\mathbb{X}_i\}$ $\Longleftrightarrow$ When evaluating the model function $\mathbb{X}(\mathbb{U})$ on a (high-resolution) uniformly, equidistantly spaced grid in $\mathbb{U}$-space, the resulting distribution of points in $\mathbb{X}$-space has (nearly) the local density behavior that the original set $\{\mathbb{X}_i\}$ has.

!

GEOMETRIC MODELING AND DENSITY – Cont'd.

- **Parametrization:** The quadratic mapping / best approximation requires one to determine a 'meaningful' parametrization. The problem is the construction of a parameter value / tuple $\mathbf{u}$ to a (feature) point / value $\mathbf{x} = \mathbf{x}(\mathbf{u})$. Ideally, when evaluating $\mathbf{u}$-space / discretizing $\mathbf{u}$-space with uniformly, equidistantly spaced $\mathbf{u}$-values, the resulting cumulative distribution of $\mathbf{x}(\mathbf{u})$-values should also be uniform, equidistant (again, in the cumulative sense).

- **Ex.:** Consider a univariate, 1D example. Given a continuous distribution of x-values, how does one construct the finite number of u_i-values for x_i-values that discretize the cumulative distribution of x-values in a uniform, equidistant manner? ("x-value-to-u-value")



Consider a simple uniform distribution $p(x)$ ("probability" or "percentage" of x).

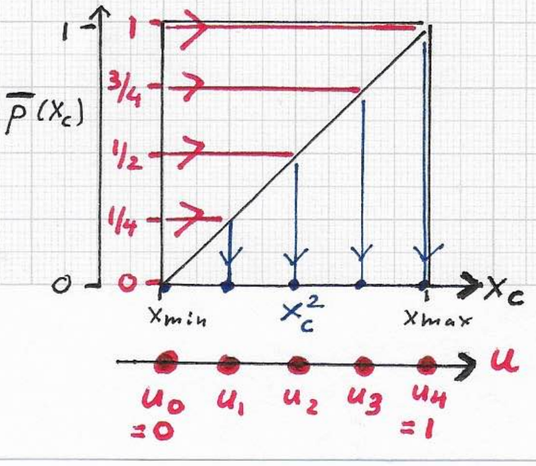
- **Here:** $x_{min} \le x \le x_{max}$, $p(x) = \frac{1}{\Delta}$, $\int_{x_{min}}^{x_{max}} p(x)\,dx = 1$.

$\Rightarrow$ cumulative distribution from x_{min} to x_c:

$$\int_{x_{min}}^{x_c} p(x)\,dx = \frac{1}{\Delta}\int_{x_{min}}^{x_c} 1\,dx = \frac{x_c - x_{min}}{\Delta} = \bar{p}(x_c).$$

$\Rightarrow$ subdivide $\bar{p}$-interval $[0,1]$ equidistantly:

$\bar{p}$-values $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ have corresponding x_c^i-values $x_{min} + {}^{i}\!/\!_{4} \cdot \Delta$, $i = 0 \ldots 4$.



$\Rightarrow$ linearly map x_c^i-values to u_i-values:

$$u_i = (x_c^i - x_{min}) / \Delta, \quad i = 0 \ldots 4.$$