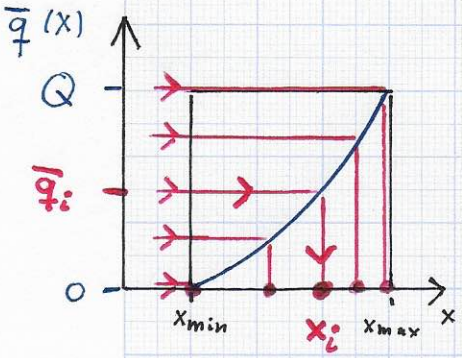
■ GEOMETRIC MODELING AND DENSITY - Cont'd.

• Best approximation & cumulative distribution :

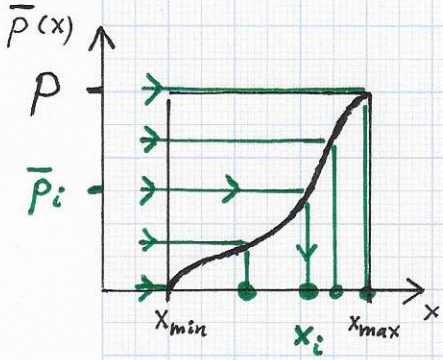
The best approximation $q(x)$ has the associated integral value $\overline{P} = \int_{x_{min}}^{x_{max}} q(x)\,dx =: Q$. The cumulative distribution of $q(x)$ is $\overline{q}(x)$, defined as $\overline{q}(x) = \int_{x_{min}}^{x} q(\tilde{x})\,d\tilde{x}$ (left figure).



generation of a synthetic set $\{x_i\}$ - based on model $q(x)$

The function $\overline{q}(x)$ can be used to generate "synthetic" discrete feature value sample values in the x-domain: For a finite set of equidistantly placed values $\overline{q}_i$ one determines the corresponding x_i-values, such that $\overline{q}(x_i) = \overline{q}_i$.

"The model function $q(x)$ is a 'good' model. ⇔ When using the cumulative functions $\overline{p}(x)$ and $\overline{q}(x)$ and sampling them with equidistantly placed $\overline{q}_i$- and $\overline{p}_i$ - values, the resulting x_i - values on the x-axis are 'nearly the same'." In other words, the simple model $q(x)$ allows one to generate discrete x-value sets with distributions that highly resemble x-value sets obtained via the correct, **original** distribution $p(x)$.

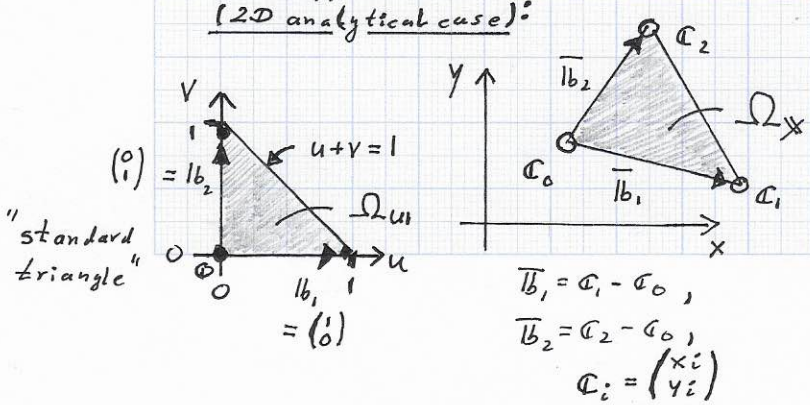


generation of a synthetic set $\{x_i\}$ - based on given original distribution

• Best approximation (2D analytical case):

In the 2D, bivariate case it is advantageous to view a 2D simplicial region Ω_X (triangle) as the image of the "standard triangle" Ω_{uv} in uv-space:

$X = X(u,v) = C_0 + u\overline{b}_1 + v\overline{b}_2$, $u, v \geq 0 \wedge u+v \leq 1$.



$\overline{b}_1 = C_1 - C_0$,
$\overline{b}_2 = C_2 - C_0$,
$C_i = \begin{pmatrix}x_i\\y_i\end{pmatrix}$

■ GEOMETRIC MODELING AND DENSITY - Cont'd.

• **Best approximation (2D case):**

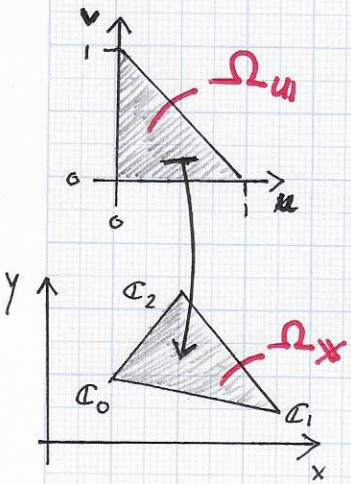
The linear mapping of the standard triangle to an arbitrary triangle in x-space can thus be written in matrix form as

$$x = Mu + c_0 = \left(\overline{b}_1, \overline{b}_2\right)\begin{pmatrix} u \\ v \end{pmatrix} + c_0 = \left(c_1 - c_0, c_2 - c_0\right)\begin{pmatrix} u \\ v \end{pmatrix} + c_0 .$$

"Change-of-Variables Theorem"

The partial derivatives of $x(u)$ with respect to u and v define the components of the Jacobian determinant of this mapping, i.e.,

$$J(u,v) = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} = \begin{vmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{vmatrix} = \left|\overline{b}_1, \overline{b}_2\right| .$$



$$\int_{\Omega_x} f(x)\,dx = J \int_{\Omega_u} f(x(u))\,du$$

This constant Jacobian determinant is needed to relate integrations performed over the u-space and x-space domain triangles, i.e., Ω_u and Ω_x:

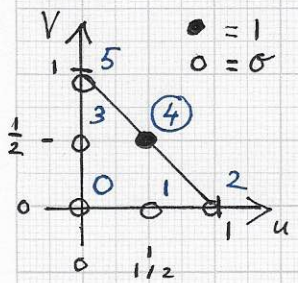
$$\int_{\Omega_x} f(x,y)\,dx\,dy = \int_{\Omega_u} f(x(u,v), y(u,v))\,J\,du\,dv .$$

Since the triangle mapping is linear $J(u,v)$ is constant, therefore written as J. This change-of-variables theorem allows one to perform integrations (inner products), necessary to compute a best approximation of a function given for the domain Ω_x, over the simple u-space domain Ω_u.

In the 2D, bivariate case, we assume that an original analytical distribution function $p(x,y)$ is given (defined over Ω_x) for which a best approximation (the model) must be computed.

■ GEOMETRIC MODELING AND DENSITY – Cont'd.

• Best approximation (2D case):



construction of quadratic basis function associated with knot (1/2, 1/2); $q_4(u,v)$ is associated with knot ④.

The distribution function $p(x,y)$ is known for Ω_X (and therefore also Ω_{UI}) in analytical form, i.e., it can be evaluated for any location in Ω_X (or Ω_{UI}). For example, $p(x,y)$ could be a simple piecewise constant function. Further, since we can use the change-of-variables theorem we only consider computations over the simple Ω_{UI} domain: We must approximate $p(u,v)$.

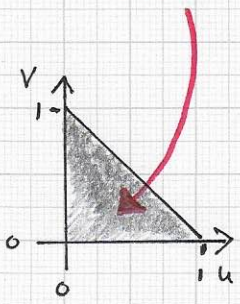
We first describe the construction of the six cardinal quadratic basis functions $q_i(u,v)$, $i=0\ldots5$, for the six knots $\left(\frac{j}{2},\frac{k}{2}\right)$, $j,k\geq0 \wedge j+k\leq2$, see left figure. Knots are indexed with a single index $(0,1,2,3,4,5)$; thus, the "cardinal property" of function $q_i(u_I)$ requires that $q_i(u_{I_\ell}) = \delta_{i,\ell}$. $\{q_i(u_I)\}$ is a basis of all quadratic polynomials defined over Ω_{UI}.

Best approximation of given $p(u_I)$:

$$q(u_I) = \sum_{i=0}^{5} c_i q_i(u_I)$$

The goal is to determine a best quadratic approximation of $p(u,v)$ using the basis $\{q_i(u,v)\}_{i=0}^{5}$.

As in the 1D case, the normal equations define the unique optimal solution (in the least-squares sense):



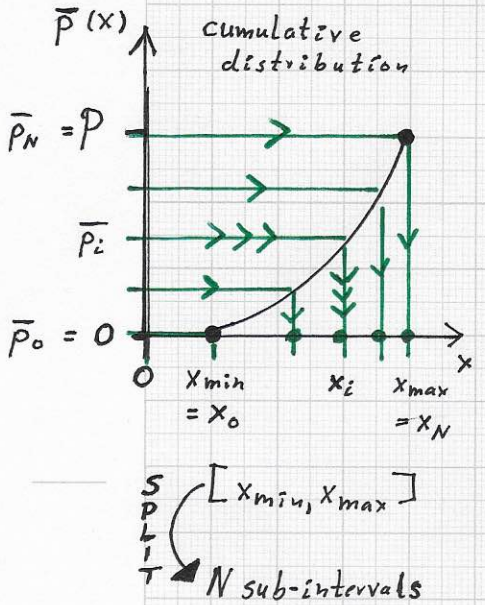
$$\langle q_i, q_j \rangle = \int_{\Omega_{UI}} q_i \cdot q_j \, du_I$$

$$\langle p, q_i \rangle = \int_{\Omega_{UI}} p \cdot q_i \, du_I$$

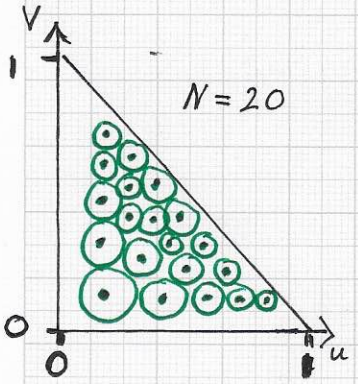
$$\begin{bmatrix} \langle q_0(u_I), q_0(u_I)\rangle & \cdots & \langle q_0(u_I), q_5(u_I)\rangle \\ \vdots & & \vdots \\ \langle q_5(u_I), q_0(u_I)\rangle & \cdots & \langle q_5(u_I), q_5(u_I)\rangle \end{bmatrix} \begin{bmatrix} c_0 \\ \\ c_5 \end{bmatrix} = \begin{bmatrix} \langle p(u_I), q_0(u_I)\rangle \\ \vdots \\ \langle p(u_I), q_5(u_I)\rangle \end{bmatrix}$$

■ GEOMETRIC MODELING AND DENSITY - Cont'd.

• Best approximation (2D case):



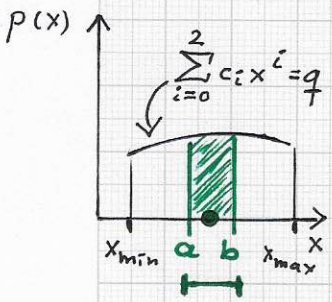
In the 1D, univariate case one starts with a distribution function ρ and computes its unique associated cumulative distribution function $\bar{\rho}$. The "cumulative range interval" $[0, P]$ is then subdivided in equal-length sub-intervals $[\bar{p}_i, \bar{p}_{i+1}]$, $i = 0 \ldots (N-1)$, with $\bar{p}_0 = 0$ and $\bar{p}_N = P$. Next, for each $\bar{p}_i$ one determines the corresponding value in the (feature space) domain Ω - leading to a synthetic feature point data set $\{x_i\}_{i=0}^{N}$. These steps must be applied to and generalized for the 2D, bivariate (and multivariate) case. For simplification, we consider u_1-space computations. The value of P is defined as $P = \int_{\Omega_{u_1}} \rho(u_1)\, du_1 = \int_{v=0}^{1-u} \int_{u=0}^{1} \rho(u,v)\, du\, dv$. The goal is to subdivide the domain Ω_{u_1} into N sub-domains/-regions; the sub-domains ("tiles") should have a distribution that "reflects" $\rho(u,v)$ and a simple geometry/representation. (Each of N tiles must satisfy $\int_{tile} \rho(u_1)\, du_1 = P/N$.) Alternatively, one can think of this domain subdivision/partioning problem as a "point packing" problem - where a point is viewed as center of a region that has the required P/N integral value.



packing method:
"bubbles" must satisfy specific conditions:
- domain coverage
- overlap/gap behavior
- distribution requirement

■ GEOMETRIC MODELING AND DENSITY - Cont'd.

• Tiling / Packing:



"tile" of '•'

$\Rightarrow \int_{tile} p(x)\,dx =$

$= \int_a^b \sum_{i=0}^{2} c_i x^i\,dx =$

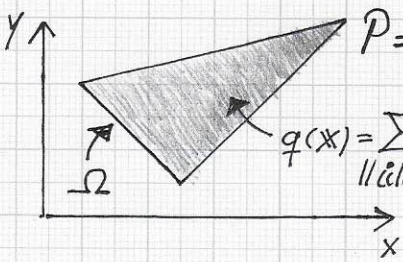
$= \sum_{i=0}^{2} c_i \frac{b^{i+1} - a^{i+1}}{i+1} =$

$= P/N$

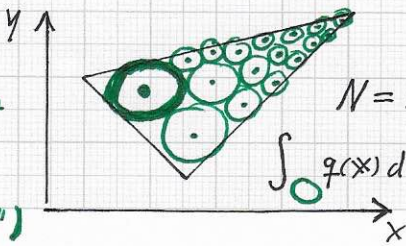
The best quadratic approximation (of an originally given LARGE discrete sample data set over a region in feature space) defines an analytical distribution model function $q(x)$. Integrating $q(x)$ over its domain yields the integral value P. When using the model $q(x)$ to generate a synthetic data set of N "samples" one must ensure that these samples have associated "tiles" that all have a value P/N as integrated distribution value (when integrating over all tiles' regions). Further, these tiles should (i) all lie inside the domain of $q(x)$; (ii) should not overlap; (iii) should have no gaps between them; and (iv) should have a total area/volume/hyper-volume that is equal to the area/volume/hyper-volume of q's domain. Thus, one must solve a tiling/packing problem.

⇒ Samples '•':

$\int_{tile(\bullet)} p(x)\,dx = P/N$



In the multivariate case where $q = q(x_1, \ldots, x_D)$, one must use an optimization method to place N samples in a domain Ω such that the tile integrals $\int_\Omega q\,dx = P/N$.



tiles and samples '•' in 2D case (tiles = "bubbles")

∴ BH

Ref.: Kenji Shimada, Physically based mesh generation: ... bubble packing, 1993.