Statistické testování algoritmů

Radek Mařík

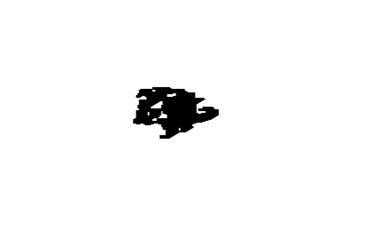
ČVUT FEL, K13133

September 6, 2011

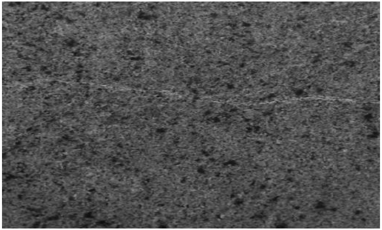
Obsah

Zpracování textury

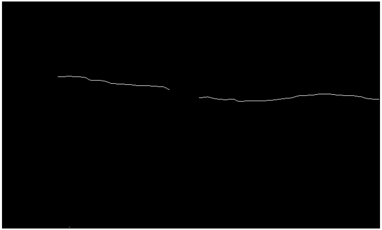
(a) Původní obraz



(b) Detekovaná skrvna

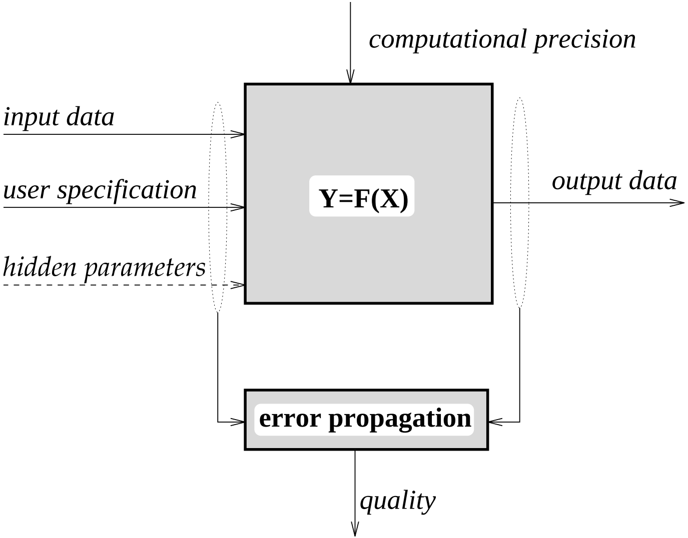


(c) Původní obraz



(d) Detekovaná prasklina

Model systému



Propagace kovarianční matice

- vstupní data: $\mathbf{X}$
- výstupní data: $\mathbf{Y}$
- explicitní vztah: $\mathbf{Y} = \mathbf{F}(\mathbf{X})$

Propagace kovarianční matice:

$$\mathbf{\Sigma}_Y = \mathbf{J}\mathbf{\Sigma}_X\mathbf{J}^T$$

kde

- $\mathbf{\Sigma}_Y$ ... kovarianční matice výstupních dat $\mathbf{Y}$,
- $\mathbf{\Sigma}_X$ ... kovarianční matice vstupních dat $\mathbf{X}$,
- $\mathbf{J}$ ... lineární operátor prvního řádu Taylorova rozvoje funkce $\mathbf{F}(\mathbf{X})$ v okolí bodu $\mathbf{X}_0$.

Propagace kovariační matice - odvození

explicitní vztah: $\underline{\mathbf{Y} = \mathbf{F}(\mathbf{X})}$

Taylorův rozvoj v okolí bodu $\mathbf{X}_0$:

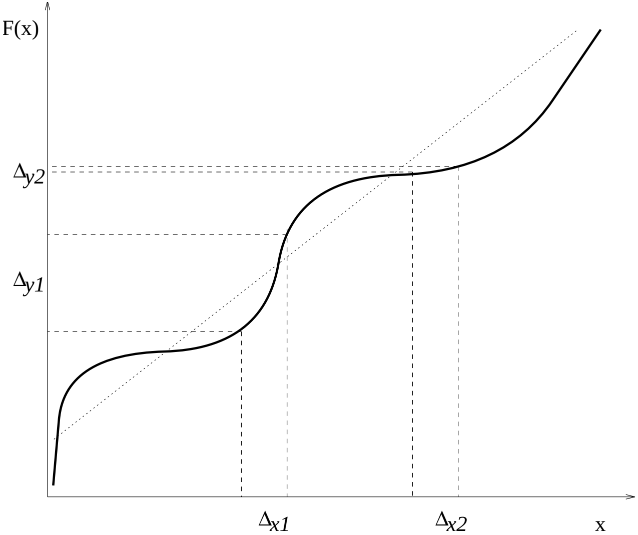
$$\mathbf{Y} = \mathbf{Y}_0 + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\mathbf{X}_0} (\mathbf{X} - \mathbf{X}_0) + \mathbf{R}_{\mathbf{Y}_0}$$

substituce:

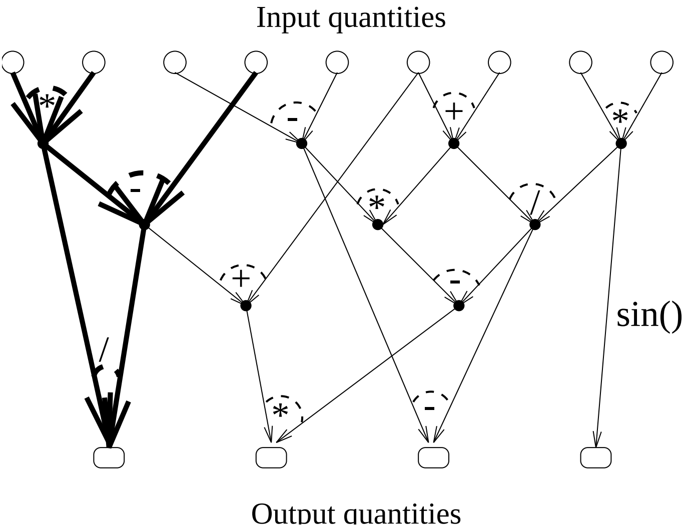
$$\begin{aligned} \delta\mathbf{X} &= \mathbf{X} - \mathbf{X}_0, \qquad \delta\mathbf{Y} = \mathbf{Y} - \mathbf{Y}_0 \\ \mathbf{J}_0 &= \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\mathbf{X}_0} \ldots \text{Jakobián} \end{aligned}$$

$$\begin{aligned} \delta\mathbf{Y} &= \mathbf{J}_0 \cdot \delta\mathbf{X} + \mathbf{R}_{\mathbf{Y}_0} \\ \delta\mathbf{Y} \cdot \delta\mathbf{Y}^T &= (\mathbf{J}_0 \cdot \delta\mathbf{X} + \mathbf{R}_{\mathbf{Y}_0})(\mathbf{J}_0 \cdot \delta\mathbf{X} + \mathbf{R}_{\mathbf{Y}_0})^T = \mathbf{J}_0 \cdot \delta\mathbf{X} \cdot \delta\mathbf{X}^T \cdot \mathbf{J}_0^T + \mathbf{R}_{\mathbf{YY}_0} \\ \underline{\mathbf{\Sigma}_{\mathbf{YY_0}}} &= 1/n\Sigma_{i=1}^{n}(\delta\mathbf{Y}_i \cdot \delta\mathbf{Y}_i^T) = 1/n\Sigma_{i=1}^{n}(\mathbf{J}_0 \cdot \delta\mathbf{X}_i \cdot \delta\mathbf{X}_i^T \cdot \mathbf{J}_0^T) = \\ &= \mathbf{J}_0 \cdot \underbrace{1/n\Sigma_{i=1}^{n}(\delta\mathbf{X}_i \cdot \delta\mathbf{X}_i^T)}_{\mathbf{\Sigma}_{\mathbf{XX_0}}} \cdot \mathbf{J}_0^T = \underline{\mathbf{J}_0 \cdot \mathbf{\Sigma}_{\mathbf{XX_0}} \cdot \mathbf{J}_0^T} \end{aligned}$$

Propagace chyby v nelineárním systému



Model datového toku



Přímé propagování chyby

- Odpadá zdlouhavé odvození chybového modelu.
- **Substituce hodnot proměnných strukturami** modelu chyby.
- “Symbolická derivace” následovaná výpočtem.
- Bodová analýza.
- Výpočet okamžité chyby.
- Lze hodnotit přesnost výpočtu.
- Předpokládá dostupnost zdrojového kódu.
- Chybový model jakéhokoliv hodnoty proměnné je vypočten pomocí chybových modelů vstupních operandů operace.
- Parametry chybových modelů:
  - dáno,
  - určené přesností měření,
  - určené předchozími výpočty.

Přímé propagování chyby - implementace

- Založeno na přetížení operátorů v objektově-orientovaném kódování.
- Přepínání mezi normálním kódem a kódem propagující chyby.

C++ normální kód

```
typedef double RealT;
```

C++ kód propagující chyby

```
class RealT  .......;
```

Příklad

```
RealT x, a=3, b=5;
x = a + b;
```

Propagace variance

Dána funkce $g(x, y)$ dvou proměnných x a y s variancemi σ_{xx}^2 a σ_{yy}^2 a kovariancí σ_{xy}.

Variance hustoty chyb hodnot funkce

$$\begin{aligned}\sigma_{gg}^2(x_0, y_0) \quad = \quad & \sigma_{xx}^2(\partial g/\partial x)^2\big|_{(x_0,y_0)} \\ & + \sigma_{yy}^2(\partial g/\partial y)^2\big|_{(x_0,y_0)} \\ & +2\,\sigma_{xy}(\partial g/\partial x)(\partial g/\partial y)\big|_{(x_0,y_0)}\end{aligned}$$

Typické zjednodušení: $\sigma_{xy} = 0$

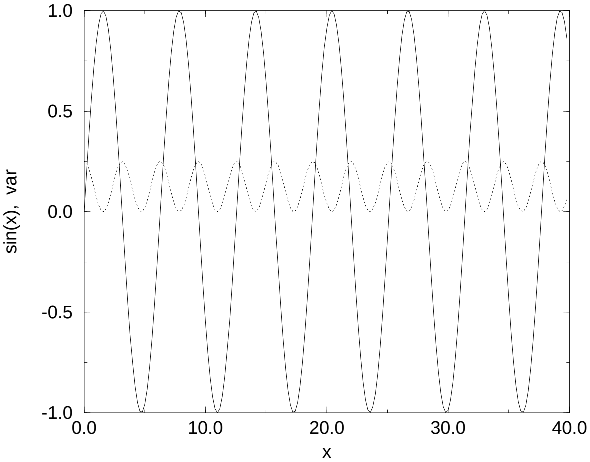
$$\sigma_{gg}^2(x_0, y_0) = \sigma_{xx}^2(\partial g/\partial x)^2\big|_{(x_0,y_0)} + \sigma_{yy}^2(\partial g/\partial y)^2\big|_{(x_0,y_0)}$$

Propagace variance

Základní aritmetické operace a standardní funkce:

Funkce	Propagace variance
$z = x + y$	$\sigma_{zz}^2 = \sigma_{xx}^2 + \sigma_{yy}^2$
$z = x - y$	$\sigma_{zz}^2 = \sigma_{xx}^2 + \sigma_{yy}^2$
$z = xy$	$\sigma_{zz}^2 = y_0^2\sigma_{xx}^2 + x_0^2\sigma_{yy}^2$
$z = x/y$	$\sigma_{zz}^2 = (\sigma_{xx}^2 + x_0^2/y_0^2\sigma_{yy}^2)/y_0^2$
$z = x^2$	$\sigma_{zz}^2 = 4x_0^2\sigma_{xx}^2$
$z = \sqrt{x}$	$\sigma_{zz}^2 = \frac{1}{4x_0}\sigma_{xx}^2$
$z = \sin x$	$\sigma_{zz}^2 = \cos^2 x_0\, \sigma_{xx}^2$
$z = \cos x$	$\sigma_{zz}^2 = \sin^2 x_0\, \sigma_{xx}^2$
$z = e^x$	$\sigma_{zz}^2 = e^{2x_0}\, \sigma_{xx}^2$
$z = \ln x$	$\sigma_{zz}^2 = 1/x_0^2\sigma_{xx}^2$
$z = \log_a x$	$\sigma_{zz}^2 = \frac{1}{(x_0 \ln a)^2}\sigma_{xx}^2$
$z = x^y$	$\sigma_{zz}^2 = (y_0^2/x_0^2\sigma_{xx}^2 + \ln^2 x_0\, \sigma_{yy}^2)\, x_0^{2y_0}$

Propagace variance - příklad



Hodnoty funkce $\sin x$ pro argument s absolutní chybou ± 0.25.

Důsledky zanedbání korelace

- hloubková data ovlivněná chybnou kalibrací a náhodným šumem,
- $\tilde{d}_i$ ... správné vzdálenosti,
- chybná kalibrace b se standardní odchylkou σ_b,
- šum n_i se společnou standardní odchylkou σ_n,
- předpoklad: kalibrace a šum jsou nezávislé se střední hodnotou šumu i odchylky kalibrace 0.

$$\begin{aligned} d_i &= \tilde{d}_i + b + n_i \\ \sigma^2_{d_i} &= \sigma^2_b + \sigma^2_n \\ \sigma_{d_1 d_2} &= \sigma^2_b \end{aligned}$$

$$\rho = \frac{\sigma_{d_1 d_2}}{\sigma^2_b + \sigma^2_n} = \frac{\sigma^2_b}{\sigma^2_b + \sigma^2_n} = \frac{1}{1 + \frac{\sigma^2_n}{\sigma^2_b}}$$

Důsledek

- Korelace je nezanedbatelná pokud odchylka je mnohem větší než přesnost standardní odchylky šumu.
- $\sigma_n = 1mm$ a $\sigma_b = 3mm$ by vedl k $\rho = 0.9$, tedy 90% korelaci.

Lineání závislost - vstupy

$$\mathbf{Y}^T = \begin{vmatrix} d_1 & d_2 \end{vmatrix}$$

$$\mathbf{X}^T = \begin{vmatrix} b & n_1 & n_2 \end{vmatrix}$$

$$\mathbf{Y} = \mathbf{F}(\mathbf{X}) = \begin{vmatrix} \tilde{d}_1 + b + n_1 & \\ \tilde{d}_2 + b & + n_2 \end{vmatrix}$$

$$\mathbf{J} = \frac{\partial \mathbf{F}}{\partial \mathbf{X}} = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix}$$

$$\mathbf{\Sigma_{XX}} = \begin{vmatrix} \sigma_b^2 & 0 & 0 \\ 0 & \sigma_n^2 & 0 \\ 0 & 0 & \sigma_n^2 \end{vmatrix}$$

Lineání závislost - propagace

$$
\begin{aligned}
\mathbf{\Sigma_{YY}} &= \mathbf{J}\cdot\mathbf{\Sigma_{XX}}\cdot\mathbf{J}^T = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} \bullet \begin{vmatrix} \sigma_b^2 & 0 & 0 \\ 0 & \sigma_n^2 & 0 \\ 0 & 0 & \sigma_n^2 \end{vmatrix} \bullet \begin{vmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{vmatrix} = \\
&= \begin{vmatrix} \sigma_b^2 & \sigma_n^2 & 0 \\ \sigma_b^2 & 0 & \sigma_n^2 \end{vmatrix} \bullet \begin{vmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} \sigma_b^2+\sigma_n^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2+\sigma_n^2 \end{vmatrix} = \\
&= (\sigma_b^2+\sigma_n^2)\begin{vmatrix} 1 & \frac{\sigma_b^2}{\sigma_b^2+\sigma_n^2} \\ \frac{\sigma_b^2}{\sigma_b^2+\sigma_n^2} & 1 \end{vmatrix} = (\sigma_b^2+\sigma_n^2)\begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix} = \\
&= \begin{vmatrix} \sigma_{d_1}^2 & \sigma_{d_1 d_2} \\ \sigma_{d_1 d_2} & \sigma_{d_2}^2 \end{vmatrix}
\end{aligned}
$$

Rozdíl korelovaných proměnných

$$
\begin{aligned}
\mathbf{Y} &= f(\mathbf{X}) = \Delta = d_2 - d_1 \\
\mathbf{X}^T &= \begin{vmatrix} d_1 & d_2 \end{vmatrix} \\
\mathbf{J} &= \frac{\partial \mathbf{F}}{\partial \mathbf{X}} = \begin{vmatrix} -1 & 1 \end{vmatrix} \\
\mathbf{\Sigma_{XX}} &= \sigma_n^2 \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix} \qquad \begin{aligned} d_1 &= \tilde{d}_1 + n_1 \\ d_2 &= \tilde{d}_2 + n_2 \end{aligned} \\
\mathbf{\Sigma_{YY}} &= \sigma_\Delta^2 = \mathbf{J} \cdot \mathbf{\Sigma_{XX}} \cdot \mathbf{J}^T = \begin{vmatrix} -1 & 1 \end{vmatrix} \bullet \sigma_n^2 \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix} \bullet \begin{vmatrix} -1 \\ 1 \end{vmatrix} = \\
&= \sigma_n^2 \begin{vmatrix} -1+\rho, & -\rho+1 \end{vmatrix} \bullet \begin{vmatrix} -1 \\ 1 \end{vmatrix} = \sigma_n^2(1-\rho+1-\rho) = \\
&= 2(1-\rho)\sigma_n^2
\end{aligned}
$$

Průměr korelovaných proměnných - vstupy

$$
\begin{aligned}
\mathbf{Y} &= f(\mathbf{X}) = d = \sum_{i=1}^{n} d_i/n \\
\mathbf{X}^T &= \begin{vmatrix} d_1 & d_2 & \cdots & d_n \end{vmatrix} \\
\mathbf{J} &= \frac{\partial \mathbf{F}}{\partial \mathbf{X}} = \begin{vmatrix} 1/n & 1/n & \cdots & 1/n \end{vmatrix} \\
\mathbf{\Sigma_{XX}} &= \sigma_n^2 \begin{vmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \ldots & \ldots & \ldots & \ldots \\ \rho & \rho & \cdots & 1 \end{vmatrix} \qquad d_i = \tilde{d}_i + n_i
\end{aligned}
$$

Průměr korelovaných proměnných - propagace

$$
\begin{aligned}
\mathbf{\Sigma_{YY}} &= \sigma_d^2 \\
&= \mathbf{J} \cdot \mathbf{\Sigma_{XX}} \cdot \mathbf{J}^T \\
&= \left| \begin{array}{cccc} 1/n & 1/n & \cdots & 1/n \end{array} \right| \bullet \sigma_n^2 \left| \begin{array}{cccc} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \ldots & \ldots & \ldots & \ldots \\ \rho & \rho & \cdots & 1 \end{array} \right| \bullet \left| \begin{array}{c} 1/n \\ 1/n \\ \cdots \\ 1/n \end{array} \right| = \\
&= \sigma_n^2 \left| \begin{array}{c} 1/n + (n-1)\rho/n, \cdots \end{array} \right| \bullet \left| \begin{array}{c} 1/n \\ 1/n \\ \cdots \\ 1/n \end{array} \right| = \\
&= \sigma_n^2 (1/n + (n-1)\rho/n) = \frac{1 + (n-1)\rho}{n} \sigma_n^2
\end{aligned}
$$

Vliv průměrů a rozdílů

- Průměrná vzdálenost

$$d = \sum_{i=1}^{n} d_i/n \qquad \sigma_d = \sqrt{\frac{1+(n-1)\rho}{n}}\sigma_n$$

- Rozdíl dvou vzdáleností

$$\Delta = d_2 - d_1 \qquad \sigma_\Delta = \sqrt{2(1-\rho)}\sigma_n$$

- Průměrování korelovaných pozorování má pouze omezený vliv, $\lim_{n\to\infty} \sigma_d = \sqrt{\rho}\sigma_n$. Např. 90% korelace omezuje zdola standardní odchylku na $0.95\sigma_n$.
- Standardní odchylka rozdílu je významně menší pro korelovaná data než pro nekorelovaná. Např. 90% korelace vede ke standardní odchylce $0.45\sigma_n$ přičemž pro nekorelovaná data je $1.4\sigma_n$.
- Testování průměrných hodnot vede k příliš optimistickým výsledkům, zatímco rozdíly vedou k pesimistickým výsledkům.

Propagace chyby Min/Max

Uzavřený interval reálných čísel nebo *Interval*

$$\begin{aligned} A &= [a_{min}, a_{max}] \\ &= \{t | a_{min} \leq t \leq a_{max}, a_{min}, a_{max} \in \mathcal{R}\} \end{aligned}$$

- $I(\mathcal{R})$ ... množina uzavřených intervalů,
- $A, B, C, \ldots, X, Y, Z$ ... intervaly z $I(\mathcal{R})$.

Nechť $* \in \{+, -, \cdot, :\}$ je binární operace na množině reálných čísel $\mathcal{R}$. Jestliže $A, B \in I(\mathcal{R})$, potom

$$A * B = \{z = a * b | a \in A, b \in B\}$$

definuje binární operace na $I(\mathcal{R})$.

Propagace chyby Min/Max - operace

Vstupní hodnoty:

$$x \in [x_{min}, x_{max}], y \in [y_{min}, y_{max}]$$

Funkce	Výstupní Min/Max intervaly
$z = x + y$	$[x_{min} + y_{min}, x_{max} + y_{max}]$
$z = x - y$	$[x_{min} - y_{max}, x_{max} - y_{min}]$
$z = x \cdot y$	$[\min\{x_{min} \cdot y_{min}, x_{min} \cdot y_{max}, x_{max} \cdot y_{min}, x_{max} \cdot y_{max}\}, \max\{x_{min} \cdot y_{min}, x_{min} \cdot y_{max}, x_{max} \cdot y_{min}, x_{max} \cdot y_{max}\}]$
$z = x/y$	$[x_{min}, x_{max}] \cdot [1/y_{max}, 1/y_{min}]$

Jestliže $g(x)$ je spojitá unární operace na $\mathcal{R}$, potom

$$g(X) = [\min_{x \in X} g(x), \max_{x \in X} g(x)]$$

definuje *unarní operarci* na $I(\mathcal{R})$.

Propagace chyby Min/Max - monotónní unární funkce

- rostoucí funkce $g_i(x)$:

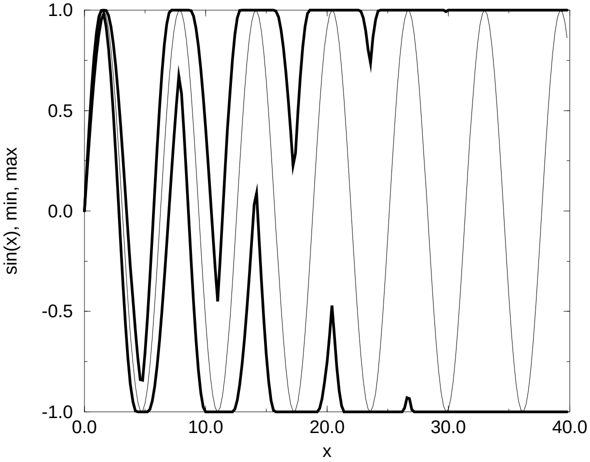
$$[x_{min}, x_{max}] \Longrightarrow [g_i(x_{min}), g_i(x_{max})]$$

- klesající funkce $g_d(x)$:

$$[x_{min}, x_{max}] \Longrightarrow [g_d(x_{max}), g_d(x_{min})]$$

Propagace chyby Min/Max - nemonotónní funkce

$\sin x$ s 10% relativní chybou argumentu



Jednoduchý příklad

Funkce

$$f(x) = x + 10$$

Její implementace:

$$f_{imp}(x) = (x \cdot x - 100)/(x - 10).$$

Testován rozsah x: $[0, 20]$ s krokem 0.01.

Jednoduchý příklad - C++ kód

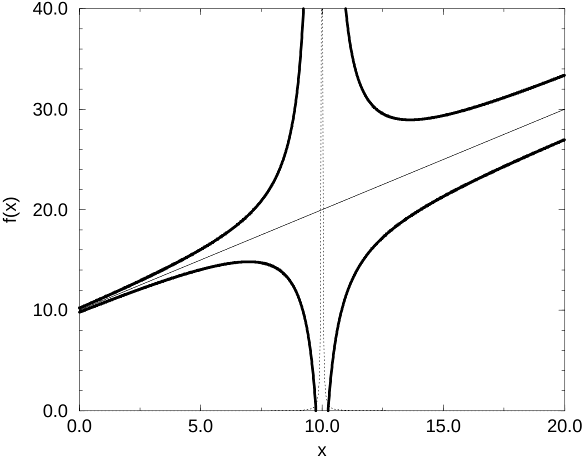
```
#include <iostream.h>
#include "StdType.hh"
int main(void)
{
  for (RealT x=0; x<=20; x+=0.01){
    if (x == 10) continue;  // to avoid division by zero
    cout << x << ' ' << (x*x -100)/(x-10) << '\n';
  }
  return 0;
}
```

StdType.hh:

- typedef double RealT;
- class RealT { .... };

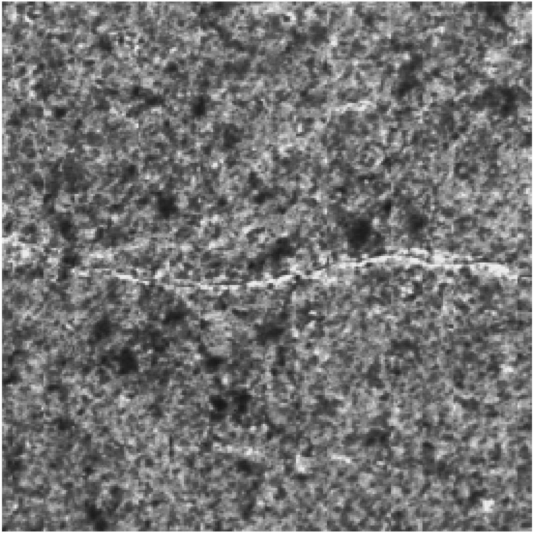
Jednoduchý příklad - výsledek

Relativní chyba 10% ve vstupních datech.

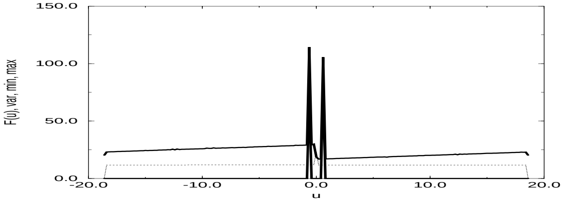


Detekce prasklin [SPK95]

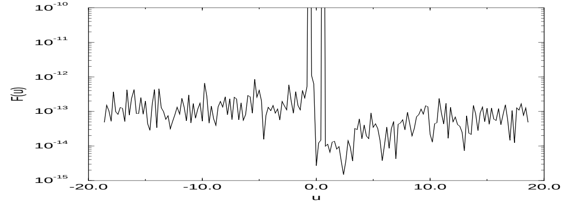
Textura žuly s horizontální prasklinou.



Výpočet DFT vzhledem k definici

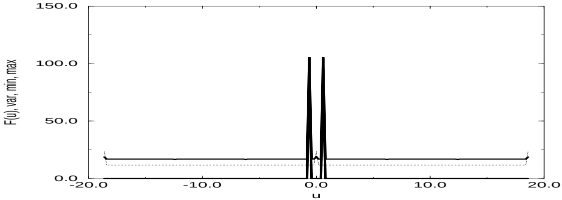


(a) 0.1% výpočetní chyba

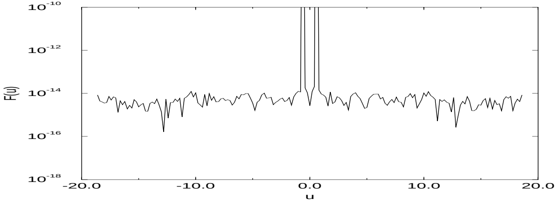


(b) přesnost procesoru - double

Výpočet DFT podle modifikace (modulo)



(a) 0.1% výpočetní chyba



(b) přesnost procesoru - double

Propagace kovarianční matice - implicitní forma [Har94]

- vstupní data: $\mathbf{X}$
- výstupní data: $\mathbf{Y}$
- vztah mezi $\mathbf{Y}$ a $\mathbf{X}$ je vyjádřen implicitní skalární funkcí $F(\mathbf{X}, \mathbf{Y})$.

Definice úlohy: dáno $\hat{\mathbf{X}} = \mathbf{X}_0 + \Delta\mathbf{X}$, určit $\hat{\mathbf{Y}} = \mathbf{Y}_0 + \Delta\mathbf{Y}$ tak, aby se minimalizovala $F(\hat{\mathbf{X}}, \hat{\mathbf{Y}})$ za předpokladu, že $\mathbf{Y}_0$ minimalizuje $F(\mathbf{X}_0, \mathbf{Y}_0)$.
Estimátor:

$$\hat{\mathbf{\Sigma}}_{\Delta\mathbf{Y}} = (\frac{\partial\mathbf{g}}{\partial\mathbf{Y}}(\hat{\mathbf{X}}, \hat{\mathbf{Y}}))^{-1} \frac{\partial\mathbf{g}}{\partial\mathbf{X}}(\hat{\mathbf{X}}, \hat{\mathbf{Y}}) \mathbf{\Sigma}_{\Delta\mathbf{X}} \frac{\partial\mathbf{g}}{\partial\mathbf{X}}(\hat{\mathbf{X}}, \hat{\mathbf{Y}})^T [(\frac{\partial\mathbf{g}}{\partial\mathbf{Y}}(\hat{\mathbf{X}}, \hat{\mathbf{Y}}))^T]^{-1}$$

kde

$$\mathbf{g}(\mathbf{X}_0, \mathbf{Y}_0) = \frac{\partial F}{\partial\mathbf{Y}}(\mathbf{X}_0, \mathbf{Y}_0)$$

Propagace kovarianční matice - odvození 1 [Har94]

- $\mathbf{X}$ ... vektor $N \times 1$ bezchybného vstupu,
- $\Delta\mathbf{X}$ ... náhodná pertubace vektoru $\mathbf{X}$,
- $\hat{\mathbf{X}} = \mathbf{X} + \Delta\mathbf{X}$ ... pozorovaný pertubovaný vstupní vektor,
- $\mathbf{\Theta}$ ... vektor $K \times 1$ parametrů,
- $\Delta\mathbf{\Theta}$ ... náhodná pertubace na $\mathbf{\Theta}$ indukovaná náhodnou pertubací $\Delta\mathbf{X}$ na $\mathbf{X}$,
- $\hat{\mathbf{\Theta}} = \mathbf{\Theta} + \Delta\mathbf{\Theta}$ ... vypočtený náhodně pertubovaný vektor parametrů,
- F ... spojitá skalární funkce, která definuje vztah mezi $\mathbf{X}$ a $\mathbf{\Theta}$ a vztah mezi $\hat{\mathbf{X}} = \mathbf{X} + \Delta\mathbf{X}$ a $\hat{\mathbf{\Theta}} = \mathbf{\Theta} + \Delta\mathbf{\Theta}$.
- Základní úloha: dáno $\hat{\mathbf{X}} = \mathbf{X} + \Delta\mathbf{X}$, určete $\hat{\mathbf{\Theta}} = \mathbf{\Theta} + \Delta\mathbf{\Theta}$ tak, aby se minimalizovala $F(\hat{\mathbf{X}}, \hat{\mathbf{\Theta}})$ za předpokladu, že $\mathbf{\Theta}$ minimalizuje $F(\mathbf{X}, \mathbf{\Theta})$.
- Pokud $\hat{\mathbf{\Theta}}$ je vypočtena explicitní funkcí $\mathbf{h}$ tak, že $\hat{\mathbf{\Theta}} = \mathbf{h}(\hat{\mathbf{X}})$, funkce F je definována takto

$$f(\mathbf{X}, \mathbf{\Theta}) = (\mathbf{\Theta} - \mathbf{h}(\mathbf{X}))^T (\mathbf{\Theta} - \mathbf{h}(\mathbf{X}))$$

Propagace kovarianční matice - odvození 2 [Har94]

- gradient $\mathbf{g}$ je vektorová funkce $K \times 1$: $\mathbf{g}(\mathbf{X}, \mathbf{\Theta}) = \frac{\partial F}{\partial \mathbf{\Theta}}(\mathbf{X}, \mathbf{\Theta})$
- Taylorovým rozvojem $\mathbf{g}$ v okolí $(\mathbf{X}, \mathbf{\Theta})$ získáme aproximaci prvého řádu:

$$\begin{aligned} \mathbf{g}^{K\times 1}(\mathbf{X} + \Delta\mathbf{X}, \mathbf{\Theta} + \Delta\mathbf{\Theta}) &= \mathbf{g}^{K\times 1}(\mathbf{X}, \mathbf{\Theta}) \\ &+ \frac{\partial \mathbf{g}^{K\times N}}{\partial \mathbf{X}}(\mathbf{X}, \mathbf{\Theta})\Delta\mathbf{X}^{N\times 1} \\ &+ \frac{\partial \mathbf{g}^{K\times K}}{\partial \mathbf{\Theta}}(\mathbf{X}, \mathbf{\Theta})\Delta\mathbf{\Theta}^{K\times 1} \end{aligned}$$

- $F(\hat{\mathbf{X}}, \hat{\mathbf{\Theta}})$ má extrém v $\hat{\mathbf{\Theta}}$, tedy $g(\hat{\mathbf{X}}, \hat{\mathbf{\Theta}}) = 0$,
  $F(\mathbf{X}, \mathbf{\Theta})$ má extrém v $\mathbf{\Theta}$, $g(\mathbf{X}, \mathbf{\Theta}) = 0$:

$$0 = \frac{\partial \mathbf{g}}{\partial \mathbf{X}}(\mathbf{X}, \mathbf{\Theta})\Delta\mathbf{X} + \frac{\partial \mathbf{g}}{\partial \mathbf{\Theta}}(\mathbf{X}, \mathbf{\Theta})\Delta\mathbf{\Theta}$$

- Relativní extrém F je relativní minimum,
  matice $\frac{\partial \mathbf{g}^{K\times K}}{\partial \mathbf{\Theta}}(\mathbf{X}, \mathbf{\Theta})$ musí být positivně definitní pro všechna $(\mathbf{X}, \mathbf{\Theta})$,
  a není proto singulární. Proto existuje $(\frac{\partial \mathbf{g}}{\partial \mathbf{\Theta}})^{-1}$.

Klasický regresní problém [Har94]

- $\mathbf{\Theta} = \mathbf{h}(\mathbf{X}) = \mathbf{J}\mathbf{X}$
- Nalezni $\mathbf{\Theta}$ pro minimum $F(\mathbf{X}, \mathbf{\Theta}) = (\mathbf{\Theta} - \mathbf{J}\mathbf{X})^T \mathbf{\Sigma}_{\mathbf{X}}^{-1} (\mathbf{\Theta} - \mathbf{J}\mathbf{X})$
- Derivace maticových forem

$$\mathbf{\Omega} = \mathbf{A}\mathbf{x} \longrightarrow \frac{\partial \mathbf{\Omega}}{\partial \mathbf{x}} = \mathbf{A}$$

$$\mathbf{\Omega} = \mathbf{x}^T \mathbf{A} \longrightarrow \frac{\partial \mathbf{\Omega}}{\partial \mathbf{x}} = \mathbf{A}^T$$

- Gradient

$$\begin{aligned} \mathbf{g}(\mathbf{X}, \mathbf{\Theta}) &= \frac{\partial F}{\partial \mathbf{\Theta}}(\mathbf{X}, \mathbf{\Theta}) = \{\mathbf{\Sigma}_{\mathbf{X}}^{-1}(\mathbf{\Theta} - \mathbf{J}\mathbf{X})\}^T + (\mathbf{\Theta} - \mathbf{J}\mathbf{X})^T \mathbf{\Sigma}_{\mathbf{X}}^{-1} \\ &= 2(\mathbf{\Theta} - \mathbf{J}\mathbf{X})^T \mathbf{\Sigma}_{\mathbf{X}}^{-1} \end{aligned}$$

- Potom

$$\frac{\partial \mathbf{g}}{\partial \mathbf{\Theta}} = 2\mathbf{\Sigma}_{\mathbf{X}}^{-1} \qquad \frac{\partial \mathbf{g}}{\partial \mathbf{X}} = -2\mathbf{\Sigma}_{\mathbf{X}}^{-1}\mathbf{J}$$

$$\mathbf{\Sigma}_{\mathbf{\Theta}} = (2\mathbf{\Sigma}_{\mathbf{X}}^{-1})^{-1}(-2\mathbf{\Sigma}_{\mathbf{X}}^{-1}\mathbf{J})\mathbf{\Sigma}_{\mathbf{X}}(-2\mathbf{\Sigma}_{\mathbf{X}}^{-1}\mathbf{J})^T(2\mathbf{\Sigma}_{\mathbf{X}}^{-1})^{T^{-1}} = \mathbf{J}\mathbf{\Sigma}_{\mathbf{X}}^{-1}\mathbf{J}^T$$

Jednoduchý příklad $y = x^2$

Explicitní propagace

$$
\begin{aligned}
\mathbf{Y} &= \mathbf{F}(\mathbf{X}) \qquad \rightsquigarrow \qquad y = x^2 \\
J &= \frac{\partial \mathbf{F}}{\partial \mathbf{X}} = 2x \\
\mathbf{\Sigma}_{y^2} &= \left|\ \sigma_y^2\ \right| = \left|\ 2x\ \right| \left|\ \sigma_x^2\ \right| \left|\ 2x\ \right| = \left|\ 4x^2\sigma_x^2\ \right|
\end{aligned}
$$

Implicitní propagace

$$
\begin{aligned}
F(\mathbf{X}, \mathbf{Y}) &\qquad \rightsquigarrow \qquad F(x; y) = (y - x^2)^2 \cdots \text{minimalizace pro 1 bod} \\
g &= \frac{\partial F}{\partial y} = 2(y - x^2) \\
\frac{\partial g}{\partial y} &= 2 \qquad \frac{\partial g}{\partial x} = -4x \\
\mathbf{\Sigma}_{yy} &= \left|\ \sigma_y^2\ \right| = \left|\ 1/2\ \right| \left|\ -4x\ \right| \left|\ \sigma_x^2\ \right| \left|\ -4x\ \right| \left|\ 1/2\ \right| = \left|\ 4x^2\sigma_x^2\ \right|
\end{aligned}
$$

Jednoduchý příklad $y = x + q$

Explicitní propagace: $q = y - x$

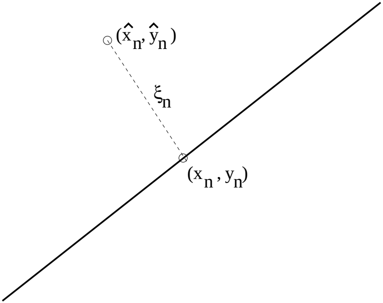
$$
\begin{aligned}
\mathbf{J} &= \left| \begin{array}{cc} \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{array} \right| = \left| \begin{array}{cc} -1 & 1 \end{array} \right| \\
\mathbf{\Sigma}_{qq} &= \left| \begin{array}{c} \sigma_q^2 \end{array} \right| = \\
&= \left| \begin{array}{cc} -1 & 1 \end{array} \right| \left| \begin{array}{cc} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{array} \right| \left| \begin{array}{c} -1 \\ 1 \end{array} \right| = \\
&= \left| \begin{array}{cc} -\sigma_x^2 + \sigma_{xy}, & -\sigma_{xy} + \sigma_y^2 \end{array} \right| \left| \begin{array}{c} -1 \\ 1 \end{array} \right| = \\
&= \left| \begin{array}{c} +\sigma_x^2 - \sigma_{xy} - \sigma_{xy} + \sigma_y^2 \end{array} \right| = \\
&= \left| \begin{array}{c} \sigma_x^2 - 2\sigma_{xy} + \sigma_y^2 \end{array} \right|
\end{aligned}
$$

Jednoduchý příklad $y = x + q$

Implicitní propagace: $y = x + q$

$$
\begin{aligned}
F(\mathbf{X},\mathbf{Y}) & \quad \rightsquigarrow F(x,y;q) = (y-x-q)^2 \cdots \text{minimalizace pro 1 bod} \\
g &= \partial F/\partial q = -2(y-x-q) \\
\partial g/\partial q &= 2 \qquad \partial g/\partial x = 2 \qquad \partial g/\partial y = -2 \\
\mathbf{\Sigma}_{qq} &= \left|\, \sigma_q^2 \,\right| = \left|\, 2 \,\right|^{-1} \left|\, 2, \ -2 \,\right| \left|\begin{array}{cc} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{array}\right| \left|\begin{array}{c} 2 \\ -2 \end{array}\right| \left|\, 2 \,\right|^{-1} = \\
&= \left|\, 1, \ -1 \,\right| \left|\begin{array}{cc} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{array}\right| \left|\begin{array}{c} 1 \\ -1 \end{array}\right| = \\
&= \left|\, \sigma_x^2 - \sigma_{xy} \quad \sigma_{xy} - \sigma_y^2 \,\right| \left|\begin{array}{c} 1 \\ -1 \end{array}\right| = \\
&= \left|\, \sigma_x^2 - 2\sigma_{xy} + \sigma_y^2 \,\right|
\end{aligned}
$$

Obecný problém aproximace přímkou - zadání [Har94]



- Nepozorované nepertubované body (x_n, y_n), $n = 1, \cdots, N$ leží na přímce $x_n \cos\theta + y_n \sin\theta - \rho = 0$
- Pozorujeme $(\hat{x}_n, \hat{y}_n)$, zašuměné instance (x_n, y_n).

Obecný problém aproximace přímkou - model [Har94]

- Model šumu:
$$\begin{pmatrix} \hat{x}_n \\ \hat{y}_n \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \xi \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$
kde ξ_n jsou nezávislé a identicky distribuované $N(0, \sigma^2)$.
- Odhad parametrů přímky $(\hat{\theta}, \hat{\rho})$ použitím metody nejmenších čtverců a kritéria:

$$F(\mathbf{X}, \mathbf{\Theta}) = \sum_{n=1}^{N} (x_n \cos\theta + y_n \sin\theta - \rho)^2$$

kde $\mathbf{X}^T = (x_1, y_1, \cdots, x_N, y_N)$ a $\mathbf{\Theta}^T = (\theta, \rho)$.

Obecná aproximace přímkou 2 [Har94]

$$
\begin{aligned}
\mu_x &= 1/N \sum_{n=1}^{N} x_n \\
\mu_y &= 1/N \sum_{n=1}^{N} y_n \\
\sigma_x^2 &= 1/N \sum_{n=1}^{N} (x_n - \mu_x)^2 = 1/N \sum_{n=1}^{N} x_n^2 - \mu_x^2 \\
\sigma_y^2 &= 1/N \sum_{n=1}^{N} (y_n - \mu_y)^2 = 1/N \sum_{n=1}^{N} y_n^2 - \mu_y^2
\end{aligned}
$$

Obecná aproximace přímkou 3 [Har94]

$$
\begin{aligned}
\sigma_{xy} &= 1/N \sum_{n=1}^{N} (x_n - \mu_x)(y_n - \mu_y) \\
&= 1/N \sum_{n=1}^{N} x_n y_n - \mu_x \mu_y \\
F(\mathbf{X}, \mathbf{\Theta}) &= \sum_{n=1}^{N} (x_n^2 \cos^2\theta + y_n^2 \sin^2\theta + \rho^2 + 2x_n \cos\theta y_n \sin\theta \\
&\quad -2\rho x_n \cos\theta - 2\rho y_n \sin\theta) \\
&= N(\sigma_x^2 + \mu_x^2)\cos^2\theta + N(\sigma_y^2 + \mu_y^2)\sin^2\theta \\
&\quad +N\rho^2 + N(\sigma_{xy} + \mu_x\mu_y)\sin 2\theta \\
&\quad -2N\rho\mu_x \cos\theta - 2N\rho\mu_y \sin\theta
\end{aligned}
$$

Aproximace přímkou 4 [Har94]

Geometrie výsledku:

- Jestliže (x, y) je bod na přímce $x\cos\theta + y\sin\theta - \rho = 0$ a k je orientovaná vzdálenost mezi (x, y) a bodem na přímce nebližším počátku, potom

$$k = \begin{cases} +\sqrt{x^2 + y^2 - \rho^2} & \text{if } y\cos\theta \geq y\sin\theta \\ -\sqrt{x^2 + y^2 - \rho^2} & \text{otherwise} \end{cases}$$

- Dále

$$\begin{aligned} x &= -k\sin\theta + \rho\cos\theta \\ y &= k\cos\theta + \rho\sin\theta \end{aligned}$$

- Nechť

$$\mu_k = 1/N\sum_{n=1}^{N} k_n \qquad \sigma_k^2 = 1/N\sum_{n=1}^{N}(k_n - \mu_k)^2$$

Aproximace přímkou - vzájemné vztahy x, y, k [Har94]

$$
\begin{aligned}
\mu_x &= \rho \cos\theta - \mu_k \sin\theta \\
\mu_y &= \rho \sin\theta + \mu_k \cos\theta \\
\sigma_x^2 &= \sigma_k^2 \sin^2\theta \\
\sigma_y^2 &= \sigma_k^2 \cos^2\theta \\
\sigma_{xy} &= -\sigma_k^2 \sin\theta \cos\theta
\end{aligned}
$$

Aproximace přímkou - interpretace [Har94]

- Po odvození a substituci

$$\mathbf{\Sigma}_{\mathbf{\Theta}} = \sigma^2/N \begin{pmatrix} \frac{1}{\sigma_k^2} & \frac{\mu_k}{\sigma_k^2} \\ \frac{\mu_k}{\sigma_k^2} & 1 + \frac{\mu_k^2}{\sigma_k^2} \end{pmatrix}$$

- Jednoduchá geometrická interpretace:
  - Souřadnicový systém přímky takový, že 0 odpovídá bodu na přímce nejbližším počátku.
  - μ_k je střední poloha bodů.
  - σ_k^2 je variance polohy bodů.
  - μ_k se chová jako momentová páka.
  - Jestliže střední poloha bodů na přímce je ve vzdálenosti $|\mu_k|$ od počátku na přímce, potom se variance odhadu ρ zvětšuje s koeficientem $\mu_k^2\sigma^2/\sigma_k^2$. Jinými slovy, odhad ρ není invariantní vzhledem k translaci souřadnicového systému.

Statistická validace algoritmu 1 [Har94]

- Testuje se, zda vypočtené odhady patří do distribuce s danou střední hodnotou a kovarianční maticí.
- Hladina významnosti α.
- Testovaná statistika $\hat{\phi}$.
- Hodnota ϕ_0 zamítnutí hypotézy.
- Postup:
  1. urči správnou odpověď pro ideální případ bez šumu,
  2. pertubuj vstupní data normálním rozložením se nulou střední hodnotou a danou kovarianční maticí,
  3. propaguj analyticky odhady kovarianční matici.

Statistická validace algoritmu 2 [Har94]

- Testovaná hypotéza: zda pozorování $\theta_1, \cdots, \theta_N$ pochází z normálního rozložení se střední hodnotou $\bar{\theta}$ a kovarianční maticí Σ.
- Existuje uniformně nejsilnější test

$$B = \sum_{n=1}^{N} (\theta_n - \bar{\theta})(\theta_n - \bar{\theta})^T$$

Definujme

$$\lambda = (e/N)^{pN/2} |B\Sigma^1|^{N/2} \times \exp(-\frac{1}{2}[tr(B\Sigma_1) + N(\bar{\theta} - \theta)^T \Sigma^1 (\bar{\theta} - \theta)])$$

- Testovaná statistika:

$$T = -2 \log \lambda$$

- T distribuováno podle $\chi^2_{p(p+1)/2+p}$ kde p je dimenze θ
- T_α: $Prob(\chi^2_{p(p+1)/2+p} \geq T_\alpha) = \alpha$

Literatura I

R.M. Haralick.
Propagating covariance in computer vision.
In *12th International Conference on Pattern Recognition (Jerusalem, Israel, 1994)*, volume I, pages 493–498, Washington, DC, 1994. IEEE Computer Society Press.

K. Y. Song, M. Petrou, and J. Kittler.
Texture crack detection.
Machine Vision and Application, 8:63–76, 1995.