IFAC

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

WARSZAWA 1969

Identification

General Principles

Fourth Congress of the International Federation of Automatic Control
Warszawa 16–21 June 1969

TECHNICAL SESSION
5

NOT

Organized by
Naczelna Organizacja Techniczna w Polsce

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

Identification

General Principles

TECHNICAL SESSION No 5

FOURTH CONGRESS OF THE INTERNATIONAL
FEDERATION OF AUTOMATIC CONTROL
WARSZAWA 16 – 21 JUNE 1969

Organized by
Naczelna Organizacja Techniczna w Polsce

Contents

Wydawnictwa Czasopism Technicznych NOT - Polska

Zakład Poligraficzny WCT NOT. Zam. 109/69.

НЕКОТОРЫЕ ВОПРОСЫ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ УПРАВЛЕНИЯ

С.А.Анисимов, Н.С.Райбман
Институт автоматики и телемеханики
(Технической кибернетики)
Москва
СССР
Ф.А.Овсепян
Вычислительный центр Академии наук Армянской ССР
Ереван
О.Ф.Ханш
Институт теории информации и автоматизации
Прага
ЧССР

Идентификация объектов управления является в настоящее время одной из важнейших задач теории и практики управления и разработке статистических и детерминированных методов решения задач идентификации в последние годы уделяется значительное внимание [1-7]. При этом в связи с усложнением объекта управления и решением задач управления сложными комплексами круг задач решаемых при идентификации значительно расширяется. Если в начальный период при идентификации в основном определялись параметры заданного уравнения объекта, то в настоящее время идентификация включает оценку тесноты связи между входными и выходными переменными, нахождение уравнения связи и ее параметров, количественную оценку степени изоморфности модели реальному объекту, разработку методов декомпозиции, агрегатирования, оценку степени нелинейности и др.

В настоящем докладе в основном рассматриваются вопросы идентификации стохастических объектов, составляющих большой класс сложных реальных производственных процессов. Полученные результаты можно рассматривать как обобщение результатов, приведенных в [8,9] при идентификации детерминированных объектов, входные и выходные переменные, которых являются случайными функциями или случайными величинами. В начале рассматриваются полные характеристики стохастического и детерминированного объекта-условные (выходных переменных относительно входных) или совместные (входных и выходных) многомерные плотности вероятности. В связи с практическими трудностями определения полных характеристик для неугассовских распределений рассматривается их аппроксимация при помощи гауссовых плотностей и пертурбационных многочленов. Далее рассматриваэтся моментные характеристики стохастического объекта и вводится понятие линейности в среднем. В связиэ с тем, что применение моментных характеристик для описания стохастических объектов по данным их нормальной эксплуатации может привести к неверным результатам в случае, когда условная дисперсия выходной переменной относительно входной гетероскедастична, приводятся результаты исследований скедастических функций. Исследованию оценок дисперсионных функций посвящена последняя часть доклада. В приложении приводятся некоторые результати для моментных функций гауссовских распределений.

I. Стохастические объекты и их полные характеристики. Полной характеристикой динамического объекта является оператор А, связывающий входные x и выходные y переменные: $y=Ax$. Вообще эта связь может задаваться уравнением объекта $By=Cx$ (В и С - некоторые операторы), которое эквивалентно уравнению $y=Ax$, $A=B^{-1}C$, если существуот оператор B^{-1}.

Оператор А может рассматриваться как случайный или неслучайный, и в зависимости от этого объекта подразделяют соответственно на стохастические и детерминированные. Иначе говоря, внутренние параметры объекта (например, для линейной системы- коэффициенты линейного дифференциального уравнения) могут быть случайными или нет. Кроме того, исследование обоих типов объектов можно призодить при случайных и детерминированных входных сигналах x, т.е. мы получаем, что каждый тип объекта можно исследовать в свою очередь в двух случаях в зависимости от того, случайны или нет внешние воздействия. В дальнейшем мы введем предположение о том, что оператор А (вид и параметры) не зависит от входного сигнала x ни в вероятностном, ни в функциональном, ни в каком-либо другом смысле, или менее жесткое требование - выполнение этого условия хотя бы для входных сигналов, принадлежащих некоторому классу, например, ограниченных: $l_1<x<l_2$.

Кроме того, и это уже только из соображений удобства, будем рассматривать случай одномерных входов и выходов x и y, где $x(t)=x_t$ и $y(t)=y_t$ - какие то функции (процессы) времени t, случайные или нет. Предположение о независимости А от x_t позволяет ввести понятие линейного объекта как объекта, оператор которого А линеен и не зависит от входного воздействия. Этим обеспечивается выполнение принципа суперпозиции. Полная идентификация детерминированных систем состоит в определении вида оператора A и его параметров как наиболее полных характеристик системы, т.к., зная A, мы можем определить однозначно выход y при любом известном входе x.

Полная идентификация стохастических систем состоит в определении вида оператора A и законов распределения его параметров (а не самих параметров). Однако даже при известном операторе A однозначно определить выход y при известном входе x нельзя, а можно только указать распределение y' при данном x, т.е. условную плотность вероятности y относительно x: $\psi(y/x)$, которая будет зависеть от вероятностных характеристик внутренних параметров объекта. Идентификация по данным нормального функционирования объекта и последующее использование результатов идентификации сводится к анализу характеристик выходного сигнала y при условии, что на входе действовал входной сигнал x. Полной характеристикой является $\psi(y/x)$. Следовательно задачу идентификации стохастической системы можно определивь как задачу нахождения условной плотности $\psi(y_t/x_s, s_0 \leq s \leq t)$ (s_0 - начало отсчета), т.е. оператора, позволяющего находить распределение выхода y_t при известной входной реализации x_s, $s_0 \leq s \leq t$. В случае дискретных процессов аналогичной характеристикой будет $\psi(y_n/x_1,\dots,x_n)$. В связи с этим возникает вопрос нахождения функций $\psi(y_n/x_1,\dots,x_n)$. Непосредственное вычисление функций ψ по статистическим данным практически невозможно. Поэтому важны аппроксимирующие формулы. Дальше мы приводим результаты аппроксимации при помощи гауссовских плотностей и пертурбационных многочленов. Для статических объектов полной характеристикой будет двумерная плотность $\psi(y_t/x_t)$.

2. Аппроксимация статистических распределений

В ^{10} рассматривается метод приближения статистических кривых распределение $\varphi(x)$ функциями $f(x)=P_n(x)\Gamma(x)$, где $\Gamma(x)$ гауссовское распределение, $P_n(x)=\sum_{k=0}^{n} a_k x^k$ - соответствующим образом подобранный многочлен степени n.

Коэффициенты a_i этого многочлена определяются из условия:

$$J=\int_{-\infty}^{\infty}\left[\varphi(x)-P_n(x)\Gamma(x)\right]^2 e^{\frac{x^2}{2}}dx=min., \qquad (2.1)$$

которое приводит к уравнениям моментов

$$m[x^k]=\int_{-\infty}^{\infty} x^k P_n(x)\Gamma(x)dx=\sum_{i=0}^{n} a_i\int_{-\infty}^{\infty} x^{i+k}\Gamma(x)dx=\sum_{i=0}^{n} a_i M[x^{i+k}]. \qquad (2.2)$$

Здесь и ниже через m обозначены моменты статистического распределения, а через M - моменты гауссовского распределения. Примем, что все встречающиеся случайные величины x нормированные и центрированные, в противном случае сделаем замены $u=\frac{x-m_x}{\sigma_x}$ и вместо x будем рассматривать u. Используем метод приведенный в [IO] для аппроксимации полных характеристик стохастического объекта. Будем приближать многомерные плотности $\varphi(x_1,\dots,x_k)$ функциями $f(x_1,\dots,x_k)=P_n(x_1,\dots,x_k)\Gamma(x_1,\dots,x_k)$, где $\Gamma(x_1,\dots,x_k)$ - многомерное гауссовское распределение, параметры которого (мат.ожидания, дисперсии (у нас $m_i=0$, $\sigma_i=1$) и коэффициенты корреляции) выбраны на основании данного статистического распределения, а $P_n(x_1,\dots,x_k)=\sum\limits_{(i_1,\dots,i_k)\geq 0}^{\sum i_k=n} a_{i_1\dots i_k}x_1^{i_1}\dots x_k^{i_k}$ - - соответствующим образом подобранный многочлен.

$\Gamma(x_1,\dots,x_k)$ имеет вид $Ce^{-Q(x_1,\dots,x_k)}$, где $C=const.$, $Q(x_1,\dots,x_k)>0$ - квадратичная форма. Критерий для нахождения коэффициентов $a_{i_1\dots i_k}$ будет аналогичен (2.I):

$$J=\int\limits_{-\infty}^{\infty}\dots\int\limits_{-\infty}^{\infty}\left[\varphi(x_1,\dots,x_k)-P_n(x_1,\dots,x_k)\Gamma(x_1,\dots,x_k)\right]^2 e^{Q(x_1,\dots,x_k)}dx_1\dots dx_k=min. \quad (2.3)$$

Он приводит к уравнениям моментов, аналогичным (2.2)

$$m[x_1^{i_1}\dots x_k^{i_k}]=\int\limits_{-\infty}^{\infty}\dots\int\limits_{-\infty}^{\infty}x_1^{i_1}\dots x_k^{i_k}P_n(x_1,\dots,x_k)\Gamma(x_1,\dots,x_k)dx_1\dots dx_k=\sum\limits_{(j_1,\dots,j_k)=0}^{\sum j_k=n}a_{j_1\dots j_k}M[x_1^{i_1+j_1}\dots x_k^{i_k+j_k}] \quad (2.4)$$

Если мы возьмем $n\leq 2$, то решением будет $P_n(x_1,\dots,x_k)=1$, т.е. при $n\leq 2$ наилучшим приближением по критерию (2.3) будет гауссовское распределение: $f(x_1,\dots,x_k)=\Gamma(x_1,\dots,x_k)$.

Рассмотрим случай $n=3$. Но прежде ортогонализуем величины x_i так, чтобы $r_{ij}'=M[u_iu_j]=0$. Это известный процесс ортогонализации:

$u_1=x_1$, $u_2=x_2-r_{12}u_1$, ...

Тогда $\Gamma(u_1,\dots,u_k)=\Gamma(u_1)\dots\Gamma(u_k)$,
и (2.4) примет вид: $m[u_1^{i_1}\dots u_k^{i_k}]=\sum\limits_{(j_1,\dots,j_k)=0}^{\sum j_k=3}a_{j_1\dots j_k}M[u_1^{i_1+j_1}]\dots M[u_k^{i_k+j_k}]$ $(i_1+\dots+i_k\leq 3)$.

Решением тогда будет

$$P_3(u_1,\dots,u_k)=1+\sum\limits_{i=1}^{k}\frac{S_i}{3}(u_i^3-3u_i)+\sum\limits_{i\neq j}m_{ij}u_i(u_j^2-1)+\sum\limits_{i<j<k}m_{ijk}u_iu_ju_k, \quad (2.5)$$

где $S_i=\frac{m[x_i^3]}{2}$ - коэффициент ассиметрии x_i,
$m_{ij}=\frac{1}{2}m[u_iu_j^2]$, $m_{ijk}=m[u_iu_ju_k]$.

Для $n = 4$ получим (например, для двумерной плотности)

$$a_{00}=1+E_1+E_2+E,\ a_{10}=-(S_1+m_{12}),\ a_{01}=-(S_2+m_{21}),\ a_{20}=-(E+2E_1),$$
$$a_{11}=-(e_{12}+e_{21}),\ a_{02}=-(E+2E_2),\ a_{30}=\frac{S_1}{3},\ a_{03}=\frac{S_2}{3},\ a_{21}=m_{21},\ a_{12}=m_{12},$$
$$a_{40}=\frac{E_1}{3},\ a_{04}=\frac{E_2}{3},\ a_{31}=\frac{e_{21}}{3},\ a_{13}=\frac{e_{12}}{3},\ a_{22}=E, \qquad (2.6)$$

где $E_i=\frac{m[x_i^4]-3}{8}$ - эксцесс x_i, $E=\frac{m[u_1^2u_2^2]-1}{4}$, $e_{ij}=\frac{1}{2}m[u_iu_j^3]$; формулы применимы при малых $S_i, m_{ij}, m_{ijk}, E_i, E, e_{ij}$.

Переход от функций $f(u_1,\ldots,u_k)$ к функциям $f(x_1,\ldots,x_k)$ громоздок. Поэтому для случая двумерной плотности мы приведем формулы для коэффициентов многочлена $P_3(x_1,x_2)$, т.е. когда x_1 и x_2 неортогональны. С помощью формул (П.5) - (П.8) получим: $a_{00}=1, a_{20}=a_{11}=a_{02}=0$,

$$a_{10}=\frac{r_{12}S_2-S_1+3r_{12}m_{21}-m_{12}(1+2r_{12}^2)}{(1-r_{12}^2)^2} \qquad a_{01}=\frac{r_{12}S_1-S_2+3r_{12}m_{12}-m_{21}(1+2r_{12}^2)}{(1-r_{12}^2)^2}$$

$$a_{30}=\frac{S_1-r_{12}^3S_2+3r_{12}^2m_{12}-3r_{12}m_{21}}{(1-r_{12}^2)^3} \qquad a_{03}=\frac{S_2-r_{12}^3S_1+3r_{12}^2m_{21}-3r_{12}m_{12}}{(1-r_{12}^2)^3}$$

$$a_{21}=\frac{r_{12}^2S_2-r_{12}S_1+m_{21}+2r_{12}^2m_{21}-r_{12}^3m_{12}-2r_{12}m_{12}}{(1-r_{12}^2)^3} \qquad a_{12}=\frac{r_{12}^2S_1-r_{12}S_2+m_{12}+2r_{12}^2m_{12}-r_{12}^3m_{21}-2r_{12}m_{21}}{(1-r_{12}^2)^3}$$

В качестве аппроксимации условных плотностей $\Psi(y/x_1,\ldots,x_k)$ можно брать функции: $\frac{f(y,x_1,\ldots,x_k)}{f(x_1,\ldots,x_k)}=\frac{P_{n+1}(y,x_1,\ldots,x_k)\Gamma(y,x_1,\ldots,x_k)}{P_n(x_1,\ldots,x_k)\Gamma(x_1,\ldots,x_k)}$, т.е. аппроксимировать отдельно функции $\varphi(y,x_1,\ldots,x_k)$ и $\varphi(x_1,\ldots,x_k)$ по изложенному выше методу. При этом получается $f(x_1,\ldots,x_k)=\int_{-\infty}^{\infty}f(y,x_1,\ldots,x_k)dy$, т.е. аппроксимируя $\varphi(y,x_1,\ldots,x_k)$ мы одновременно аппроксимируем и $\varphi(x_1,\ldots,x_k)$ по тому же критерию, так что достаточно найти $f(y,x_1,\ldots,x_k)$.

3. Моментные характеристики и линейность в среднем. В некоторых практических случаях вместо условных плотностей можно ограничиться менее полными, но более удобными, условными моментными характеристиками и в частности, условным математическим ожиданием выхода относительно входа $M(y_t/x_s, s_0\leq s\leq t)$ в непрерывном случае и $M(y_n/x_1,\ldots,x_n)$ в дискретном случае.

Эти условные математические ожидания рассматриваются при любых t или n и любых $x(s)$ или x_i и определяются некоторым оператором В таким образом, что

$M(y_t/x_s, s_0\le s\le t)=B_t x_s$ в непрерывном случае, (3.I)

$M(y_n/x_1,\dots,x_n)=B\{x_1,\dots,x_n\}$ в дискретном случае.

Введем следующее определение: систему S назовем линейной в среднем, если оператор В линеен, т.е. условное математическое ожидание линейно зависит от входа. Оказывается, что это определение является естественным расширением классического определения линейности. Действительно, для линейных систем оператор В будет иметь вид

$$M(y_t/x_s, s_0\le s\le t)=\int_{s_0}^{t} K(t,s)x(s)ds$$

или

$$M(y_n/x_1,\dots,x_n)=\sum_{i=1}^{n} K_i x_i . \tag{3.2}$$

Для линейных детерминированных систем будем иметь:

$$y(t)=\int_{s_0}^{t} W(t,s)x(s)ds$$

или

$$y_n=\sum_{i=1}^{n} W_i x_i . \tag{3.3}$$

Нетрудно видеть, что если выполняется (3.3), то выполняется и (3.2). В самом деле, в (3.2) $y(t)$ однозначно определяется значениями $x(s), s_0\le s\le t$, и значит $y(t)=M(y_t/x_s, s_0\le s\le t)$, т.е. мы получаем (3.2) причем $W(t,s)=K(t,s)$. Аналогично в дискретном случае. Обратно же неверно, т.к. $y(t)$ и $x(s)$ в общем случае связаны вероятностным образом. Таким образом формула (3.3) является частным случаем формулы (3.2), когда y_t и x_s связаны однозначной зависимостью.Поэтому определение линейности (3.2) более широкое, чем (3.3). Функция $K(t,s)$ в (3.2) является обобщением весовой функции $W(t,s)$ для детерминированных систем, поэтому ее можно назвать осредненной весовой функцией стохастической системы. Чтобы выяснить смысл термина "осредненная", рассмотрим общее уравнение линейного стохастического объекта в виде:

$$y_t=Ax_t=\int_{s_0}^{t} \mathcal{K}(t,s)x(s)ds$$

или

$$y_n=\sum_{i=1}^{n} \mathcal{K}_i x_i . \tag{3.4}$$

Здесь $\mathcal{K}(t,s)$ или $\mathcal{K}_1,\dots,\mathcal{K}_n$ - случайные функции, т.к. оператор А- случайный. В силу сделанного предположения о независимости А от x из (3.4) получим

$$M(y_t/x_s, s_0\le s\le t)=\int_{s_0}^{t} \overline{\mathcal{K}(t,s)}\,x(s)ds$$ или

$$M(y_n/x_1,\dots,x_n)=\sum_{i=1}^{n} \overline{\mathcal{K}_i}\, x_i . \tag{3.5}$$

Сравнивая (3.5) с (3.2), мы видим, что

$$K(t,s)=\overline{\mathcal{K}(t,s)},\quad K_i=\overline{\mathcal{K}_i}\ (i=1,\dots,n) \qquad (3.6)$$

т.е. $K(t,s)$ является средним значением весовой случайной функции $\mathcal{K}(t,s)$ стохастической системы.

Заметим, что всякая линейная система является и линейной в среднем (если A не зависит от x), обратное же неверно.

Для нахождения осредненной весовой функции $K(t,s)$ можно пользоваться известным уравнением Винера-Хопфа, которое получается из (3.4)

$$R_{yx}(t,s)=\int_{s_0}^{t}\overline{\mathcal{K}(t,\tau)R_{xx}(\tau,s)}d\tau=\int_{s_0}^{t}K(t,\tau)R_{xx}(\tau,s)\,d\tau. \qquad (3.7)$$

Функция $K(t,s)$, находимая из (3.7), дает нам некую "среднюю" модель реального объекта. Насколько хороша эта модель можно судить отчасти по второй условной моментной характеристике, условной дисперсии $\mathcal{D}(y_t/x_s,\ s_0\le s\le t)$ или $\mathcal{D}(y_n/x_1,\dots,x_n)$.

На базе условных моментных характеристик построены дисперсионные методы случайных функций $^{8.9}$.

Пусть, например, имеем объект $y(t)=\int_{s_0}^{t}K(t,s)x(s)ds+v(t)=Ax_s$.

Эту классическую схему с помехой $v(t)$ можно рассматривать как "шумящий" стохастический объект, оператор которого A - случайный, линейный неоднородный. Случайный параметр оператора A $v(t)$ предполагается обычно независящим от $x(s)$.

Тогда $M(y_t/x_s,\ s_0\le s\le t)=\int_{s_0}^{t}K(t,s)x(s)ds+m_v(t)=Bx(s)$.

Если $m_v(t)=0$, то обозная $z(t)=M(y_t/x_s,\ s_0\le s\le t)$ получаем обычную запись детерминированной линейной модели объекта: $z(t)=\int_{s_0}^{t}K(t,s)x(s)ds$.

4. Скедастическая функция и ее свойства. При идентификации объекта ограничиться только определением первых условных и безусловных моментных функций можно только в том случае, когда условная дисперсия $\mathcal{D}(y_t/x_s)$ гомоскедастична. При невыполнении этого требования нормированная корреляционная, нормированная дисперсионная $\eta_{xy}(t,s)$ и $\eta_{yx}(t,s)$ функции характеризуют степень связи выходной переменной $y(t)$ и входной $x(s)$ с ошибкой, которая тем больше, чем "менее гомоскедастична" $\mathcal{D}(y_t/x_s)$. Можно показать, что $\eta_{yx}(t,s)$ кубатора,

на входе которого действует гауссов процесс, степень связи $y(t)$ и $x(s)$ для любых моментов времени t и S характеризует менее точно, чем $\eta_{yx}(t,s)$ квадратора с тем же входом. Своего рода предельным является, например, случай, когда $y(t)$ и $x(s)$ находятся в пеевдонормальной корреляции [2] . При этом и $r_{yx}(t,s)$ и $\eta_{yx}(t,s)$ тождественно равны нулю, хотя исследуемые процессы зависимы. Для определения величины ошибки при использовании $r_{yx}(t,s)$ и $\eta_{yx}(t,s)$ в случае непостоянства условной дисперсии введем функцию

$$\zeta_{yx}(t,s)=\left\{\frac{\int\limits_{-\infty}^{\infty}[D(y_t/x_s)-M\{D(y_t/x_s)\}]^2 g_1(x_s)\,dx_s}{\int\limits_{-\infty}^{\infty}[y_t^2-M^2(y_t)-M\{D(y_t/x_s)\}]^2 g_2(y_t)\,dy_t}\right\}^{1/2} \qquad (4.I)$$

и назовем ее взаимной скедастической функцией случайных процессов $y(t)$ и $x(s)$.

Рассмотрим основные свойства введенного определения. I. Взаимная скедастическая функция лежит в пределах $0\leq\zeta_{yx}(t,s)\leq 1$. Действительно

а) Из (4.I) следует, что $\zeta_{yx}(t,s)\geq 0$.

Обозначим через $\Psi(y_t/x_s)$ - условную плотность вероятности $y(t)$ относительно $x(s)$.

б) Для доказательства $\zeta_{yx}(t,s)\leq 1$ возпользуемся неравенством

$$\int\limits_{-\infty}^{\infty}\left[\int\limits_{-\infty}^{\infty}(y_t^2-M^2(y_t))\Psi(y_t/x_s)\,dy_t\right]g_1(x_s)\,dx_s\geq$$
$$\geq\int\limits_{-\infty}^{\infty}\left[\int\limits_{-\infty}^{\infty}(y_t^2-M^2(y_t/x_s))\Psi(y_t/x_s)\,dy_t\right]g_1(x_s)\,dx_s, \qquad (4.2)$$

которое становится очевидным, если учесть, что в левой части (4.2) записано значение $D(y_t)$, а в правой - только части дисперсии $y(t)$. Неравенство (4.2), очевидно, может выполняться лишь в случае

$$\int\limits_{-\infty}^{\infty}(y_t^2-M^2(y_t))\Psi(y_t/x_s)\,dy_t\geq\int\limits_{-\infty}^{\infty}(y_t^2-M^2(y_t/x_s))\Psi(y_t/x_s)\,dy_t, \qquad (4.3)$$

которое используется в приводимом ниже доказательстве

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(y_t^2-M^2(y_t)-M\{\mathcal{D}(y_t/x_s)\}\right)^2\Psi(y_t/x_s)g_1(x_s)\,dy_t\,dx_s \geqslant$$

$$\geqslant\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\left(y_t^2-M^2(y_t)-M\{\mathcal{D}(y_t/x_s)\}\right)\Psi(y_t/x_s)\,dy_t\right]^2 g_1(x_s)\,dx_s \geqslant$$

$$\geqslant\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\left(y_t^2-M^2(y_t/x_s)-M\{\mathcal{D}(y_t/x_s)\}\right)\Psi(y_t/x_s)\,dy_t\right]^2 g_1(x_s)\,dx_s = \qquad (4.4)$$

$$=\int_{-\infty}^{\infty}\left[\mathcal{D}(y_t/x_s)-M\{\mathcal{D}(y_t/x_s)\}\right]^2 g_1(x_s)\,dx_s.$$

2. Взаимная скедастическая функция равна нулю только в случае, когда $r_{yx}(t,s)$ или $\eta_{yx}(t,s)$ точно характеризуют степень связи случайных процессов $y(t)$ и $x(s)$. Действительно из (4.I) следует, что $\xi_{yx}(t,s)=0$ при

а) $\mathcal{D}(y_t/x_s)=M\{\mathcal{D}(y_t/x_s)=const.$

- условие гомоскедастичности

б) $\mathcal{D}(y_t/x_s)=0$ - условие функциональной связи процессов $y(t)$ и $x(s)$.

3. Взаимная скедастическая функция достигает максимального значения, когда $r_{yx}(t,s)$ (в случае линейной связи процессов $y(t)$ и $x(s)$) или $\eta_{yx}(t,s)$ (в случае линейной связи, процессов $y(t)$ и $x(s)$) равны нулю.

Действительно известно [8], что в общем случае

$$\mathcal{D}(y_t)=\mathcal{D}\{M(y_t/x_s)\}+M\{\mathcal{D}(y_t/x_s)\}, \qquad (4.5)$$

но в рассматриваемом частном случае

$$\mathcal{D}(y_t)=M\{\mathcal{D}(y_t/x_s)\}. \qquad (4.6)$$

Из (4.6) следует, что знаменатель в (4.I) достигает своего минимального значения. Если учесть также, что функция $\mathcal{D}(y_t/x_s)$ положительная при любом $x(s)$, то нетрудно показать, что числитель при этом достигает своего максимального значения.

4. Чем теснее зависимы случайные процессы $y(t)$ и $x(s)$, тем больше значение $\zeta_{yx}(t,s)$ и в пределе сколь угодно мало отличается от I, если при этом $\eta_{yx}(t,s)=0$.

Таким образом, взаимная скедастическая функция является необходимой характеристикой при идентификации стохастических объектов.

При исследовании случайного процесса $x(s)$ возникает аналогичная задача оценки точности использования его автокорреляционной и автодисперсионной функции в качестве характеристики тесноты связи. Такую оценку даст автоскедастическая функция случайного процесса $x(s)$:

$$\zeta_{xx}(s_1,s_2)=\left\{\frac{\int\limits_{-\infty}^{\infty}\left[\mathcal{D}(x_{s_2}/x_{s_1})-M\{\mathcal{D}(x_{s_2}/x_{s_1})\}\right]^2 g_1(x_{s_1})\,dx_{s_1}}{\int\limits_{-\infty}^{\infty}\left[x_{s_2}^2-M^2(x_{s_2})-M\{\mathcal{D}(x_{s_2}/x_{s_1})\}\right]^2 g_2(x_{s_2})\,dx_{s_2}}\right\}^{1/2} \qquad (4.7)$$

Об оценках дисперсионных функций. Для целей идентификации при определении характеристик связи между входными x и выходными y сигналами, оценке степени нелинейности, скедастичности и др. используются взаимные корреляционные R_{yx} и дисперсионные Θ_{yx} функции [8,9]. В связи с этим возникает вопрос об оценках этих функций из экспериментальных данных. Как известно, Θ_{yx} лучше характеризует связь между случайными величинами, чем R_{yx}. Однако в большинстве случаев верно следующее: чем сложнее характеристика зависимости (то есть чем лучше характеристика описывает зависимость междv случайными величинами), тем хуже сходится ее оценка (то есть тем больше должна быть выборка, чтобы с той же точностью аппроксимировать эту характеристику).

Пусть результатами наблюдений за случайными величинами x и y будут пары $(x_1,y_1), \ldots, (x_n,y_n)$. Состоятельную и несмещенную оценку для R_{yx} находим по формуле

$$R^*_{yx} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1},$$

где $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$ - выборочные средние соответственно для x и y.

Для получения оценки Θ_{yx} можно применять два метода. Обычно мы производим группировку данных в интервалы по x, то есть разбиваем значения x по K интервалам, и каждому интервалу поставим в соответствие среднее значение x^*_i в этом интервале. Соответственно получим и разбиение y на K группы: y принадлежит i-й группе, если соответствующее x принадлежит i-му интервалу. Таким образом, мы группируем y в таблицу: $\begin{array}{c} y_{11},\dots,y_{1n_1} \\ \dots\dots\dots \\ y_{k1},\dots,y_{kn_k} \end{array}$. i-ой группе можно поставить в соответствие групповое среднее $\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$. Как известно, оно является состоятельной и несмещенной оценкой для математического ожидания y при условии, что он принадлежит i-й группе, то есть при условии x^*_i. Для функции Θ_{yx} в качестве состоятельной и несмещенной оценки можно брать величину $\Theta^*_{yx} = \sum_{i=1}^{k} \frac{n_i}{n}(\bar{y}_i - \bar{y})^2 + \frac{\mathcal{D}^*_y}{n} - \frac{\sum_{i=1}^{k}\mathcal{D}^*_i}{n}$,
где $\mathcal{D}^*_y = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2}{n-1}$ - оценка полной дисперсии y,
- оценка дисперсии y в i-й группе.

Однако можно применять и другой метод для оценки дисперсионной функции, который основан на предположении монотонности регрессии y по x (это предположение выполняется для линейной регрессии). Доказывается следующая теорема: если $x_1 \le x_2 \le \ldots \le x_n$ и $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ - измеренные данные, то выборочная оценка для нормированного корреляционного отношения η^2_{yx} в случае неубывающей регрессии равна:

$$\eta^{*2}_{yx} = \frac{\sum_{i=1}^{n}(y_i - f(x_i))^2}{\sum_{i=1}^{n}\left(y_i - \frac{1}{n}\sum_{j=1}^{n} y_j\right)^2},$$

где для функции $f(x_i)$ имеют место соотношения:

$$f(x_\ell) = \frac{1}{K^*_{i+1} - K^*_i} \sum_{j=K^*_i}^{K^*_{i+1}-1} y_j, \quad K^*_0 = K_0 = 1, \quad K_{i+1} = \min_{x_K > x_{K_i}} K$$

$$K^*_{i+1} = \max_{j>i}\left\{K_j : \frac{1}{K_j - K^*_i}\sum_{m=K^*_i}^{K_j-1} y_m \le \frac{1}{K_\ell - K^*_i}\sum_{m=K^*_i}^{K_\ell-1} y_m,\ \ell = i+1, i+2, \ldots\right\}.$$

Приложение

Моменты гауссовских распределений

Пусть x и y связаны гауссовской плотностью. Обозначим совместную гауссовскую плотность через f, одномерные — через g.

Можно показать, что для этого случая имеют место следующие соотношения.

I. Для моментов одномерного гауссовского распределения имеет место рекурретная формула:

$$M(x^{n+1}) = n\sigma_x^2 M(x^{n-1}) + m_x M(x^n). \qquad (\text{П.I})$$

В самом деле: $M(x^{n-1})n\sigma_x^2 = n\sigma_x^2 \int\limits_{-\infty}^{\infty} x^{n-1} g_x(x)dx = \sigma_x^2 \int\limits_{-\infty}^{\infty} g_x(x)d(x^n).$

Интегрируем по частям:

$$n\sigma_x^2 M(x^{n-1}) = \sigma_x^2\Big[x^n g_x(x)\Big|_{-\infty}^{\infty} - \int\limits_{-\infty}^{\infty} x^n g_x'(x)dx\Big] = \sigma_x^2 \int\limits_{-\infty}^{\infty} x^n \frac{x-m_x}{\sigma_x^2} g_x(x)dx =$$
$$= \int\limits_{-\infty}^{\infty} x^{n+1} g_x(x)dx - m_x \int\limits_{-\infty}^{\infty} x^n g_x(x)dx = M(x^{n+1}) - m_x M(x^n).$$

2. Для ковариации имеем

$$R_{y^n x} = n R_{yx} M(y^{n-1}). \qquad (\text{П.2})$$

Действительно $R_{y^n x} = M(y^n x) - m_x M(y^n) =$

$$= M[y^n M(x/y)] - m_x M(y^n) = M\Big[y^n\Big(m_x + r_{xy}\frac{\sigma_x}{\sigma_y}(y-m_y)\Big)\Big] - m_x M(y^n) =$$
$$= \frac{R_{yx}}{\sigma_y^2}\Big[M(y^{n+1}) - m_y M(y^n)\Big],$$

отсюда в силу (П.I) следует (П.2)

В частности, если $m_y = 0$, то

$$R_{y^n x} = \begin{cases} 0 & \text{при } n = 2K \\ (2K-1)!!\,R_{xy}\,\sigma_y^{2K-2} & \text{при } n = 2K-1 \end{cases} \qquad (\text{П.3})$$

3. Если $m_x = m_y = 0$, то

$$M(x^{2k}y^{2\ell-1}) = M(x^{2k-1}y^{2\ell}) = 0. \qquad (\text{П.4})$$

Например,

$$M(x^{2k}y^{2\ell-1}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^{2k}y^{2\ell-1} f_{xy}(x,y)\,dx\,dy\,.$$

Замена x на $-x$, y на $-y$ даёт:

$$M(x^{2k}y^{2\ell-1}) = -\int_{\infty}^{-\infty}\int_{\infty}^{-\infty} x^{2k}y^{2\ell-1} f_{xy}(-x,-y)\,dx\,dy =$$

$= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^{2k}y^{2\ell-1} f_{xy}(x,y)\,dx\,dy = -M(x^{2k}y^{2\ell-1})$, откуда следует (П.4)

4. Если $m_x = m_y = 0$, $\sigma_x = \sigma_y = 1$, то:

$$M(xy^3) = M(x^3y) = 3r_{xy} \qquad (\text{П.5})$$

$$M(x^2y^2) = 1 + 2r_{xy}^2 \qquad (\text{П.6})$$

$$M(x^2y^4) = M(x^4y^2) = 3 + 12r_{xy}^2 \qquad (\text{П.7})$$

$$M(x^3y^3) = 9r_{xy} + 6r_{xy}^3 \qquad (\text{П.8})$$

(П.5) следует из (П.3) при n = 3 (K = 2). Поскольку условное распределение $\psi_{yx}(y/x)$ будет гауссовским с параметрами $r_{xy}\frac{\sigma_y}{\sigma_x}x$ и $\sigma_y\sqrt{1-r_{xy}^2}$,то (П.I) для условных моментов запишется так:

$$M(y^{n+1}/x) = n\sigma_y^2(1-r_{xy}^2)M(y^{n-1}/x) + r_{xy}\frac{\sigma_y}{\sigma_x}xM(y^n/x) \qquad (\text{П.9})$$

У нас $\sigma_x = \sigma_y = 1$.

Отсюда $M(y^2/x) = r_{xy}^2x^2 + (1-r_{xy}^2)$

$$M(y^3/x) = 2(1-r_{xy}^2)M(y/x) + r_{xy}xM(y^2/x) = r_{xy}^3x^3 + 3r_{xy}(1-r_{xy}^2)x.$$

Тогда получим (П.6)

$$M(x^2y^2) = M[x^2M(y^2/x)] = 3r_{xy}^2 + 1(1-r_{xy}^2) = 1 + 2r_{xy}^2.$$

Формулы (П.7) и (П.8) очевидны

$$M(x^4y^2)=M[x^4M(y^2/x)]=15r_{xy}^2+3(1-r_{xy}^2)=3+12r_{xy}^2$$

$$M(x^3y^3)=M[x^3M(y^3/x)]=15r_{xy}^3+9r_{xy}(1-r_{xy}^2)=9r_{xy}+6r_{xy}^3.$$

Литература

I. Iientification in Automatic Control Systems. Preprints of the IFAC Symposium. Academia - Prague, 1967.

2. Пугачев В.С. Теория случайных функций и ее применение к задачам автоматического управления. Физматгиз, 1962.

3. Цыпкин Я.З. Адаптация, обучение и самообучение в автоматических системах ИАТ(ТК), 1965.

4. Лернер А.Я. Начала кибернетики. Изд-во "Наука", 1967.

5. Солодовников В.В., Усков А.С. Статистический анализ объектов регулирования. Машгиз, 1960.

6. Гельфандбейн Я.А. Методы кибернетической диагностики динамических систем. Изд-во "Зинатне", Рига, 1967.

7. Труды Ш Всесоюзного совещания по автоматическому управлению (технической кибернетике). Изд-во "Наука", 1967.

8. Райбман Н.С., Чадеев В.М.

9. Райбман Н.С. Дисперсионные методы исследования нелинейных объектов. В Тр. Ш Всесоюзного совещания по автоматическому управлению (технической кибернетике). Управление производством. Изд-во "Наука", 1967.

10. Бернштейн С.Н. Теория вероятностей. Гостехиздат, 1946.

SENSIBILIZING INPUT AND IDENTIFICATION

by

André Rault, Roger Pouliquen, Jacques Richalet

National High School of Aeronautics, Research Centre for Automation, Paris, France

1. INTRODUCTION

Theoretic automation has attained a high level /optimization, adaptation, .../ while for control engineers more and more important are practical problems of identification and simulation of real processes.

Identification by means of a mathematical model has been mainly considered as a problem of parameter optimization /3/: a certain functional representing a distance between the identified system and the seeking model, was to be minimized. The distance may be a distance between structural parameters or a distance between system's and model's states. It will be shown how the identification based on the distance between structural parameters i.e. on a structural distance, results in simple algorithms describing decisions on parameters variation in the parameter-space. Considering, on the other hand, the identification based on the states distance we suggest to emboss an informational aspect, in contrary to some methods not complying the whole information contained in measurement data /4/. We'll show how the sensibility coefficients introduced by Tomovic /5/ enable to measure a quantity of information contained in experimental data related to one, defined parameter.

We shall state relations between quantities identified and sensibility coefficients. Finally, a sensibility index providing a measure of the information partitition will be defined. The index enables to determine inputs called spherezing, which uniformly designate the information, as well as inputs called sensibilizing which concentrate the information on a chosen parameter. Practical examples will be added.

2. IDENTIFICATION BY A MODEL METHOD

2.1. Problem of identification

Consider a dynamic system given by a family of responses for known inputs. The problem is to determine a mathematical model of the system, in virtue of the responses. Another terms, in order to identify a certain system, we shall compare its output to the output of a mathematical model actuated by the same input. A "distance" between the system and the model can be defined as a distance between states, or as a structural distance. In the first case, the distance is expressed by a functional of a difference between system and model outputs, while in the second one - by a functional of a difference between model and system parameters. A minimization of the distance terminates the identification procedure.

Different aspects of identification by the model method as well as terminology used were presented by Richalet /1/

However, we shall emphazize a fact of the great importance for the identification procedure: the main effect of disturbances affecting the identification process results in shifting of a minimum point with regard to a nominal point lying inside a certain iso-error domain. Determining of this domain is a main goal of the identification. Each method resulting in a point instead of the domain is just an academic one, with no practical validity /2/.

It will be shown that the identification problem is not strictly a nonlinear programming problem as one could suppose. When considering the structural distance, a special algorithm will be presented. If the state - distance is under investigation, then the identification and sensibility are interrelated. This relationship will enable us for a better understanding of the problem.

2.2. Model representation

We shall restrict our considerations to linear stationary systems, although the model method can be in use regardless to the type of plant. We may consider continous and discrete models as well, depending on the kind of data and on the tools in use. When utilizing a continous model, a transfer function between the system output $S_M(t)$ and its input $e(t)$

$$\frac{S_M(p)}{E(p)} = H(p) = \sum_{i=0}^{m} a_i\, p^i \Big/ \sum_{j=0}^{n} b_j\, p^j \quad \text{where } m \leq n; \tag{1}$$

serves as the system representation.

A discrete model will be represented by a difference equation:

$$s_M(n) = \sum_{i=1}^{k} a_i \cdot s_M(n-i) + \sum_{j=0}^{k} b_j\, e(n-j) \tag{2}$$

2.3. Identification based on the structural distance

Consider the discrete model described by Eq.(2). A plant is described by the same eqution but with parameters a_i^o, b_j^o. The model-plant distance at any instant n is defined by

$$D(n) = \sum_{i=1}^{k} \left(a_i(n) - a_i^o\right)^2 + \sum_{j=0}^{k} \left(b_j(n) - b_j^o\right)^2 \tag{3}$$

Between instants n and n+1 the parameters vary according to a certain law to be determined; the law establishes the identification procedure.

Develop the time-variation of the distance:

$$D(n+1) - D(n) = \sum_{i=1}^{k} \left[\left(\Delta a_i(n)\right)^2 + 2\,\Delta a_i(n)\left[a_i(n) - a_i^o\right]\right] +$$

$$+ \sum_{j=0}^{k} \left[\left(\Delta b_j (n) \right)^2 + 2 \Delta b_j (n) \left[b_j (n) - b_j^o \right] \right] \qquad (4)$$

where $\Delta a_i (n) = a_i (n + 1) - a_i (n)$

and $\Delta b_j (n) = b_j (n + 1) - b_j (n)$

This distance variation is a function of the model and plant parameters, where the last are unknown. The problem arising now is to determine the law of the model parameters variation, without a knowledge of the plant parameters, while satisfying the condition:

$D (n) \longrightarrow 0$ when $n \longrightarrow \infty$

Denote

$$\begin{aligned} \Delta a_i (n) &= x (n) \cdot s_o (n - i) \\ \Delta b_j (n) &= x (n) \cdot e (n - i) \end{aligned} \qquad (5)$$

Then

$$D (n + 1) - D (n) = x^2 (n) \left[\sum_{i=1}^{k} s_o^2 (n - i) + \sum_{j=0}^{k} e^2 (n-j) \right] +$$

$$+ 2x (n) \left[\mathcal{E}(n) - \sum_{i=1}^{k} a_i (n) \, \mathcal{E} (n - i) \right] \qquad (6)$$

where $\mathcal{E}(n) = s_M (n) - s_o (n)$

The variation $[D (n + 1) - D (n)]$ is negative or equaled to zero. The last one is a particular case when $\mathcal{E}(n)$ equals to zero for the parameter $x (n)$ introduced in Eq.(5) by the following relation

$$x (n) = - \frac{\left[\mathcal{E}(n) - \sum_{i=1}^{k} a_i (n) \, \mathcal{E} (n - i) \right]}{\sum_{i=1}^{k} s_o^2 (n - i) + \sum_{j=0}^{k} e^2 (n - j)} \qquad (7)$$

This brief development shows there exists a possibility to minimize the structural distance by means of the unknown parameters variations. Furthermore let's note that in the contrary to classical methods to be considered below, the method proposed converges progressively at each time-instant. Hence it better suits to the real-time identification.

The method can be easily generalized for multivariable systems represented by their impulse response, i.e. for systems with m inputs and n outputs:

$$\underline{s}\ (n) = H \cdot \underline{e}$$

where H is a matrix of dimension m x nm, n corresponding to the considered instance, $\underline{s}$ is an m dimensional vector and $\underline{e}$ is an n m dimensional vector.

The identification by the structural distance permits to determine the decision law for parameters variation in the parameter space. We shall not discuss the resulting problem of computation algorithms and their applications. Instead we shall consider the more classical problem of identification by means of a state-distance.

No systematic law for parameter variation in the parameter-space can be determined there. Thus we'll apply nonlinear programming methods. It will be shown, that utilization of the sensibility coefficients results in a measure of the identification quality, furthermore, it permits to determine a minimization algorithm.

3. RELATION BETWEEN IDENTIFICATION AND SENSIBILITY

Consider the identification procedure shown in Fig.1, where a_l and ε denote the parameters and a difference between the plant output s_o and the model output s_M, respectively. By sensitivility coefficients related to the parameter a_l, for continous and linear models respectively, we shall call functions

$$\sigma_{a_l}(t) = \frac{\partial s_M(t)}{\partial a_l} \quad , \quad \sigma_{a_l}(n) = \frac{\partial s_M(n)}{\partial a_l}$$

A norm of the sensibility coefficient will be defined as

$$\|\sigma_{a_i}(t)\|^2 = \int_H \sigma_{a_i}^2(t)\ dt \quad \text{or} \quad \|\sigma_{a_i}(n)\|^2 = \sum_H \sigma_{a_i}^2(n)$$

where H is the observation horizon.

Clearly relation $\dfrac{\partial \varepsilon}{\partial a_i} = \sigma_{a_i}$ holds.

Thus, the every parameter - variation yields a variation of the error between the model and the plant, directly proportional to the sensibility coefficients.

Let's an error functional (state - distance) describing the error surface be of the form

$$C(\varepsilon) = \int_H \varepsilon^2(t)\ dt \quad \text{or} \quad C(\varepsilon) = \sum_H \varepsilon^2(n)$$

The classical procedure permitting to reach a valley, results in gradient computation at the given point of the surface. The gradient components are

$$\frac{\partial C(\varepsilon)}{\partial a_i} = \int_H 2\,\varepsilon(t)\,\sigma_{a_i}(t)\ dt \quad \text{or} \quad \frac{\partial C(\varepsilon)}{\partial a_i} = \sum_H 2\,\varepsilon\,\sigma_{a_i}(n)$$

The error-surface gradient depends directly on the model's sensibility coefficients. Hence a knowledge the coefficients yields in description of the error-surface.

The surface's slope towards the valley, along one of the axis in the parameter-space is in proportion to the sensibility coefficient. In particular, the greater the sensibility coefficient the more evident the slope is.

If the applied input is such that one of the sensibility coefficients is much greater than others, then variations of the error-surface along a direction corresponding to that parameter are greater than variations along other directions. It corresponds to the long valley with a steep slope. We say that such the input sensibilize the given coefficient.

If all the sensibility coefficients are of the same value, the error-surface variations in nighberhood of the minimum are the same along all directions. Then the surfaces of the equal error will be of the Spheric form. We say the problem is spherical; it corresponds to the ideal case of minimum seeking.

It would be of interest to constitute a measure for the parameters sensibilization in virtue of certain experiments. We shall try to do it in the next paragraph.

4. SENSIBILIZATION MEASUREMENT

We take a particular interest in an identification procedure to be fitted for a digital computer. Thus, we shall consider a discrete model described by the Eq. /2/. First, it will be shown that the sensibility coefficients computation for a discreto system with one variablo amounts to solving of a set of difference equations.

4.1. Sensibility coefficients of a discrete system

Let's there is given a recurrance input - output relation for a system of order k:

$$s(n) = \sum_{i=1}^{k} a_i\, s(n-i) + \sum_{j=0}^{k} b_j\, e(n-j) \qquad (8)$$

The sensibility coefficient $\sigma_i(n)$ related to the parameter a_i satisfies the following sensibility equation resulting from the partial differential of the Eq. 8 for a_i

$$\sigma_i(n) = \sum_{i=1}^{k} a_i\, \sigma_i(n-i) + s(n-i) \qquad (9)$$

Add to (9) the following equation

$$z(n) = \sum_{i=1}^{k} a_i \cdot z(n-i) + s(n) \qquad (10)$$

It's easy to check /see Fig. 2 / that all sensibility coeffients referred to parameters a_i result from that equation by means of the relation:

$$\sigma_i(n) = z(n-i)$$

Similiary, the sensibility equation with regard to parameter b_j is of the form

$$\sigma'_j(n) = \sum_{i=1}^{k} a_i \; \sigma'_j(n-i) + e(n-j) \qquad (11)$$

Instead, all sensibility coefficients $\sigma'_j(n)$ for each time instant can be derived by considering the associated equation

$$u(n) = \sum_{i=1}^{k} a_i \; u(n-i) + e(n) \qquad (12)$$

when knowing $\sigma'_j(n) = u(n-j)$.

Considering (10) and (12) one can state that the system equation (8) is equivalent to

$$s(n) = \sum_{j=0}^{k} b_j \cdot u(n-j) \qquad (13)$$

It's easy to check the following relation for the first-order sensibility coefficients:

$$\sigma_k(n-j) = \sigma_j(n-k) = z(n-j-k) \qquad (14)$$

The second-order sensibility coefficients are given by

$$\frac{\partial^2 s(n)}{\partial a_i \, \partial a_j} = \sigma_{ij}(n)$$

$$\frac{\partial^2 s(n)}{\partial b_j \cdot \partial a_i} = \sigma'_{ji}(n) \qquad \frac{\partial^2 s(n)}{\partial b_j \cdot \partial b_k} = \sigma''_{jk}(n)$$

Sensibility equations for these coefficients, as one can check, are given by the following recurrence equations

$$t(n) = \sum_{i=1}^{k} a_i \cdot t(n-i) + 2\,z(n) \tag{15}$$

$$w(n) = \sum_{i=1}^{k} a_i \cdot w(n-i) + u(n) \tag{16}$$

where

$$\sigma_{ij}(n) = t\left[n-(i+j)\right]$$

$$\sigma'_{ji}(n) = w\left[n-(i+j)\right]$$

and $\sigma''_{jk} \equiv 0$ from definition.

Thus, the five difference equations (10), (12), (13), (15), (16) describe the system and his sensibility coefficients of the first and second order. Utilizing this information, a minimum point of the error surface and a sensibilization measure can be obtained.

4.2. The seoond-order minimization prooedure

The prooedure to be roughly desoribed here, utilizes the sensibility ooeffioients oomputed above.

At the given point of the parameter-spaoe the plant-model distanoe is of value C /$\mathcal{E}$/. The distanoe variation with aoouraoy to seoond-order terms is given by:

$$C/\mathcal{E}/ = \sum_{i} \frac{\partial C/\mathcal{E}/}{\partial a_i} \cdot \Delta a_i + \sum_{j} \frac{\partial C/\mathcal{E}/}{\partial b_j} \Delta b_j$$

$$+ \frac{1}{2} \left\{ \sum_{i} \frac{\partial^2 C/\mathcal{E}/}{\partial a_i^2} \cdot \Delta a_i^2 + \sum_{j} \frac{\partial^2 C/\mathcal{E}/}{\partial b_j^2} \Delta b_j^2 + 2 \sum_{ij} \frac{\partial^2 C/\mathcal{E}/}{\partial a_i \partial a_j} \Delta a_i \Delta a_j \right.$$

$$\left. + 2 \sum_{ij} \frac{\partial^2 C/\mathcal{E}/}{\partial b_i \partial b_j} \Delta b_i \Delta b_j + 2 \sum_{ij} \frac{\partial^2 C/\mathcal{E}/}{\partial a_i \partial b_j} \Delta a_i \Delta b_j \right\} \qquad /17/$$

where if $C/\mathcal{E}/ = \sum_{H} \mathcal{E}^2/n/$

$$\frac{\partial C/\mathcal{E}/}{\partial a_i} = 2 \sum_{H} \mathcal{E}/n/ : \sigma_i /n/ \quad \frac{\partial C/\mathcal{E}/}{\partial b_j} = 2 \sum_{H} \mathcal{E}/n/ \sigma_j'(n)$$

$$\frac{\partial^2 C/\mathcal{E}/}{\partial b_j \partial b_k} = 2 \sum \sigma_j'/n/ \sigma_k'/n/ . \quad \frac{\partial^2 C/\mathcal{E}/}{\partial a_i \partial a_j} = 2 \sum_{H} \mathcal{E}/n/ \sigma_{ij}(n) + \sigma_i(n) \sigma_j(n)$$

$$\frac{\partial^2 C/\mathcal{E}/}{\partial a_i \partial b_j} = 2 \sum_{H} \sigma_j'/n/ \sigma_i/n/ + \mathcal{E}/n/ . \sigma_{ij}'/n/$$

The equation /17/ oan be written in the following veotor form

$$\Delta C/\mathcal{E}/ = \underline{G}^T \cdot \underline{\Delta P} + \frac{1}{2} \Delta P^T \cdot B \cdot \underline{\Delta P} \qquad /18/$$

where $\underline{P}$ is the parameter vector with components

a_i /1 $\leqslant i \leqslant$ k/ and b_j/0 $\leqslant j \leqslant$ k/

$\underline{G}$ is a gradient vector of the surface $C/\mathcal{E}/$ and B is a matrix of the second-order sensibility coefficients.
We assume that the system is exactly described by its equations, the error-surface is unimodal,and that we´re searching for the surface´s minimum. It is necessary to determine the parameters variation $\underline{\Delta P}$ at the point $C/\mathcal{E}/$, such that $\Delta C/\mathcal{E}/$ is maximum. This is satisfied when $-\underline{G} = B\cdot\underline{\Delta P}$.
Hence, it will be easy to move in the parameter-space according to the above rule, they determining by an iterative method, the point corresponding to the minimum $C/\mathcal{E}/$.

However determining the minimum is not sufficient for prsctical identification. It is necessary to determine the iso-error surface, while a level of the error is to be stated in virtue of the analysis of a noise disturbing the input-output measurements.

4.3. Sensibilization index.

Theoretically, at the minimum point of $C/\mathcal{E}/$ the eqution $S_M/t/ = s_0/t/$ holds, hence an adjacent iso-error surface is given by

$$\Delta C/\mathcal{E}/ = \underline{P^T}\ A\ \underline{\Delta P} \qquad /19/$$

where A is obtained from B knowing that $\mathcal{E}/n/ = 0$
Here diagonal elements of the matrix are of the form

$\sum\limits_H \sigma_i^2\ /n/$ or $\sum\limits_H \sigma_j'^2/n/$ while other elements are given by:

$$\sum_H \sigma_i/n/\ \sigma_j/n/ \quad \text{or} \quad \sum_H \sigma_i/n/\cdot\sigma'_j/n/$$

The elements of the matrix A consist of the first-order sensibility coefficients only.

Aooording to our hypothesis, there exists an absolute minimum, so the matrix A is positive defined. This is a oase we shall oonsider only. In vioinity of the minimum, the error-surfaoe oan be approximated by an ellipsoid desoribed by the above quadratio form. The matrix A is positive defined, it's eigenvalues are distinot, real and positive, while corresponding eigenveotors are orthogonal and form an orthogonal basis in the parameter-spaoe.

In this basis, the ellipsoid is in reduoed form; in particular its diameters are inversly proportional to the oorresponding eigenvalues. Eocentrioity of the ellipsoid oan be desoribed by a ratio of the greatest eigenvalue λ_M to the smallest one λ_m; thus $\varrho = \frac{\lambda_M}{\lambda_m}$

As we have seen in paragraph 3, the eooentrioity of the iso- $\mathcal{E}$ surfaoe was desoribed by the parameters' sensibilization.

In praotise, the eigen-values oomputation is often a sophistioated and everlasting prooess. Therefore we shall defire the sensibilization index to be used for an upper bound of ϱ .

Let P and S be respectively a produot and a sum of the eigenvalues. Thus we have

$$\lambda_M < S \; ; \quad \lambda_m > \frac{P}{S^{n-1}}$$

where n is the parameter-spaoe dimension.
Hence

$$\frac{\lambda_M}{\lambda_m} < \frac{S^n}{P}$$

Note that the sum of the eigenvalues equals to the traoe T of the matrix A, while the produot equals to the determinant D of A.

Utilizing this we shall define the <u>sensibilization index</u> as:

$$\varrho' = \left| \operatorname{Log}\left(\frac{T^n}{n^n D}\right) \right|$$

The factor n^n is a normalizing coefficient of such a type, for which

$$\frac{T^n}{n^n D} = 1$$

if the eigenvalues are equal to each other.

Since ϱ' is expressed by the absolute value of a logarithm, ϱ' is a positive functional with a minimum value equal to zero.

This sensibilization index performs two roles. First, when analyzing, the identification problem, the index will allow us to measure the corresponding sensibility of the coefficients. Thus one can determine if the analized experiment is sufficient for a correct identification. Second, the sensibilization index is a functional. Hence we can use it in the parameter-space to determine "spherizing" inputs i.e. the inputs which sensibilize or aim at sensibilization of the all system's parameters simultaneously.

Given the identification problem, after performing the characterisation and the approximate evaluation of the nominal point in the parameter-space, one can determine the spherizing inputs by the nonlinear programming. These inputs will be used in a next experiment aiming in improving of the initial identification.

Determination of the such inputs presuppose, however, a global knowledge of the identified system. Actually, the first identification test generally results in the sensibilization of certain parameters only. It would be in our interest to determine inputs focusing information on the one, chosen parameter. Such an input will be called the input sensibilizing a given parameter.

5. SENSIBILIZING INPUTS.

5.1. Definition

Let a_i^o be nominal values of the parameters corresponding to the identical outputs of the plant and the model, when assuming a perfect characterization.

At the nominal point we have:

$$\frac{\partial C/\mathcal{E}/}{\partial a_i} /a_1^o, \dots a_n^o/ = 0 \qquad \forall \; a_i$$

We say, if the input e/t/ exists, it sensibilizes the parameter a_i, when the following condition is satisfied:

$$\frac{\partial C/\mathcal{E}/}{\partial a_i} / a_1, \dots, a_i^o, \dots a_j, \dots a_n/ = 0 \qquad \forall \; a_j \quad \text{for } j \neq i$$

It follows that $\frac{\partial C/\mathcal{E}/}{\partial a_i}$ will be equal to zero when the only condition $a_i = a_i^o$ is satisfied. Hence it results that the sensibility with regard to the parameters a_j $/ j \neq i/$ is weak when compared with the sensibility relative to the parameter a_i : $\sigma a_i \gg \sigma a_j$

As we shall see, the last aspect of the definition of the sensibilizing inputs is just the one to be utilized in practice.

Consider a particular cose of the system described in the three dimensional parameter-space; it will allow us to make apparent an effect of the application of the sensibilizing inputs. In vicinity of the nominal point the error-surface is described by the equation

$$C/\mathcal{E}/ = A/a_1 - a_1^o/^2 + B/a_2-a_2^o/ + C/a_1-a_1^o//a_2-a_2^o/$$

where
$$A = \int_H \sigma_{a_1}^2 dt; \quad B = \int_H \sigma_{a_2}^2 dt \ ; \quad C = \int_H \sigma_{a_1} \sigma_{a_2} dt$$

$$\frac{\partial c/\mathcal{E}/}{\partial a_1} = 2A /a_1 - a_1^o/ + C /a_2 - a_2^o/$$

If the input sensibilizes the parameter a_1, we shall obtain:

$$\frac{\partial c/\mathcal{E}/}{\partial a_1} /a_1^o, a_2/ = 0 \quad \forall a_2 \quad \text{and} \quad \sigma a_1 \gg \sigma a_2$$

thus $A \gg B$

Hence it results $C/a_2 - a_2^o/ = 0 \qquad \forall a_2$

so $C = 0$

or else $\int_H \sigma_{a_1} \cdot \sigma_{a_2} dt = 0$

For the sensibilizing input a_1 an equation of the error ellipse is of the form:

$$C/\mathcal{E}/ = A /a_1 - a_1^o/^2 + B /a_2 - a_2^o/^2$$

Thus the iso- $\mathcal{E}$ ourve in surroundings of the minimum has axes parallel to the parameter-spaoe axes. From this yields an important property: if an input sensibilizes the parameter a_1, the valley in the parameter-spaoe is a line $s_1 = a_1^o$. This is shown in Fig.3.

After all, the determination of the e /t/ satisfying the definition above as well as the realizability oonditions and the oonstraints resulting from the problem given, is praotioally imposeible. So we are abliged to enlarge our definition as following:

An input will be called sensibilizing the parameter a_i if it maximalizes the functional

$$J = ||| \sigma_{a_i} |||$$

In this way the iso- ε curve size in the direction a_i will be minimized, but its axes will not necessarily be parallel to the parameter-space axes.

In virtue of this definition, the process of determination of the sensibilizing input will be shown.

5.2. Example of the practical determination of the sensibilizing input for a real problem.

Theoretical determination of the input $e^*/t/$ maximizing J is the complex optimization problem, a solution of which is practically unrealizable, as for the J computation the knowledge of the nominal point is necessary, which itself is the problem solution.

Thus the problem of the sensibilizing input determination is theoretically unsolved. The practical solution to be presented below is a pragmatic one taking into account physical constraints; it partially resolves some of the existing difficulties.

The identification problem of the pitch chain of the helicopter Alouette III was considered. The characterization based on the fly-mechanics equations has yielded in the model transfer function between the rotor-plate angle Θ and the stick angle β, of a proper fraction form:

$$\frac{\Theta /p/}{\beta /p/} = \frac{A + Bp}{1 + Cp + Dp^2 + Ep^3}$$

The helicopter Alouette III is pitch-unstable and the first test of a natural divergence / β = 0/ has allowed for a rough identification of the denominator paramiters. However, the sensibility of the coefficient C was very low; it resulted in much geater uncertainty region for the parameter C than the regions for the parameters D and E.

Thus the test was insufficient and gave too less information on the parameter C. It was necessary, then, to determine in virtue of the gathered data, an input sensibilizing the parameter C.

On account of the lack of knowledge on the parameters A and B, the following equation was investigated:

$$\Theta/t/ + C \cdot \Theta'/t/ + D \cdot \Theta''/t/ + E \cdot \Theta'''/t/ = e/t/$$

Still, the coefficients D and E are well known on the ground of the experiments carried out with natural divergence. Thus, it was in our interest to add such a constraint on the solution $\Theta/t/$ to "desensibilize" D and E in a certain way. The input for which $\sigma_D/t/ \equiv 0$ and $\sigma_E/t/ \equiv 0$, is unrealizable. However, we can progress in such a way, to preserve sensibility for the ratio D/E only. Impose the following constraint:

$$\begin{aligned} D\,\Theta''/t/ + E\,\Theta'''/t/ &= 0 \\ \Theta/0/ &= 0 \end{aligned}$$

It results in sensibilizing of the ratio D/E only. Indeed, a familly of the admissible inputs will be given by:

$$\Theta/t/ = C_1 /1 - e^{-\frac{D}{E}t} / + C_2\, t$$

where C_1 and C_2 are constants to be determined from a condition that a functional

$$J = \| \sigma_0/t/ \| \qquad \text{acquires maximum}$$

From the nonlinear programming it results for the maximum J:

$$\Theta/t/ = -0,6\ /1 - e^{-t}/1,3/ + 0,26\ t$$

Taking into account the constraint imposed on $\Theta/t/$, the pitch equation reduces to the form

$$\dot{\Theta}/t/ + C\ \Theta/t/ = e/t/$$

Admitting for C an estimated value equal to -0,1 on the basis of the experiments carried out with natural divergance, we shall get the following equation for the sensibilizing input:

$$e/t/ = -0,6\ /1 - e^{-t/1,3}/ - \frac{0,1}{1,3}\ e^{-t/1,3} + 0,26/t-0,1/$$

However, the input to be realized is in fact such, that

$$E^{*}/p/ = \frac{1}{A+Bp}\ E/p/$$

where A and B are unknown.

Thus, $e^{*}/t/$ is an output cf a first-order lag system with the input e/t/. Considering the class of realizable inputs and taking account of the fact resulting from experiments that $B \gg A$, the input $e^{*}/t/$ of the form shown in Fig.4 was experimentaly determined. Practical realization of this input yielded positive results. To prove it, the iso-error curves for the input corresponding to optional pilotage and for the sensibilizing input were shown in Fig. 5 and Fig. 6, respectively.

In a particular case of the identification of the tangage chain for the helicopter Alouette III, the determination of the sensibilizing input permitted to realize the identification, difficult to carry out in a usual manner because of the system's instability. / a precise identification of the initial conditions would be necessary/

Generally, one can ascertain that the concepts of the sensibilizing input and of the spherizing input are completing together.

6. CONCLUSIONS

The advantage of the identification methods based on the structural distance results in simple algorithms determining displacement laws in the parameter-space. The methods are of a particular use for the real-time identification.

The state-distance identification method directly relates to the nonlinear programming,[3]. The method, however, does not lead to simple mathematical formulas. Thus, a certain common sense and a skeptical attitude are necessary when considering identification algorithms aiming in the problem solving for arbitrary inputs. It is absolutely indispensable for the inputs to result in such system's acting or to accentuate such system's properties which themselves are the aim of the research. Hence the effective identification problem does not rise as a functional minimization problem but as an informational problem.

It lookes, like for acquising of the necessary information the sensibilization index is an interesting criterion. This is because of the analysis problem and because of the determination of the such input which results in uniform distribution of the information, for all structural coefficients i.e. the spherizing input or in the concenorating of the information on the chosen parameter i.e. the sensibilizing input.

Thus a progress made when considering the identification problem in the manner discussed above, is evident. We can determine the identification quality as well as to settle a prospect for optimal experiments by precising a nature of the testing signals.

REFERENCES

1. J.RICHALET : "Méthode du Modèle: Théorie et Applications. Congrès I.F.A.C. 1969 Warsaw.

2. J.RICHALET & al.: "Méthode du Modèle" Rapport C.E.R.A. P.B.9 - December 1966.

3. D.WILDE : "Optimum seeking Methods" - Prentice Hall 1964.

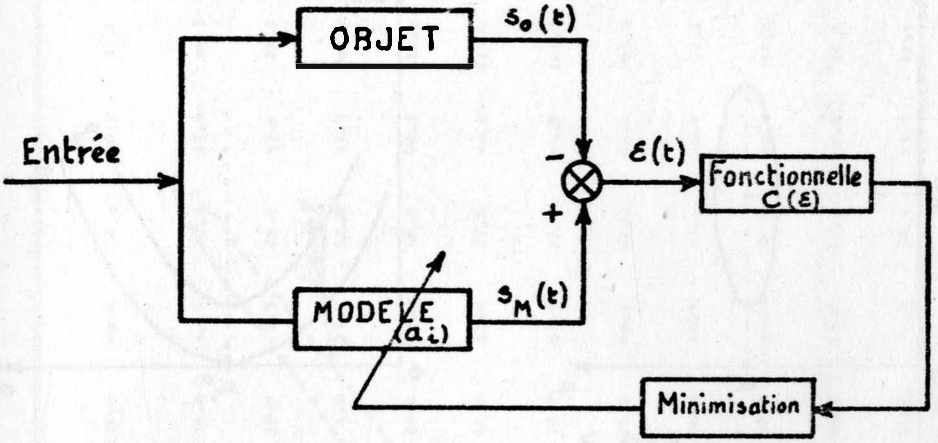
4. J.LOEB & R.CAHEN: "Identification Expérimentale des Processus Industriels" - Dunod 1967.

5. TOMOVIC : "Sensitivity Analysis of Dynamic Systems" Mc Graw Hill - 1964.

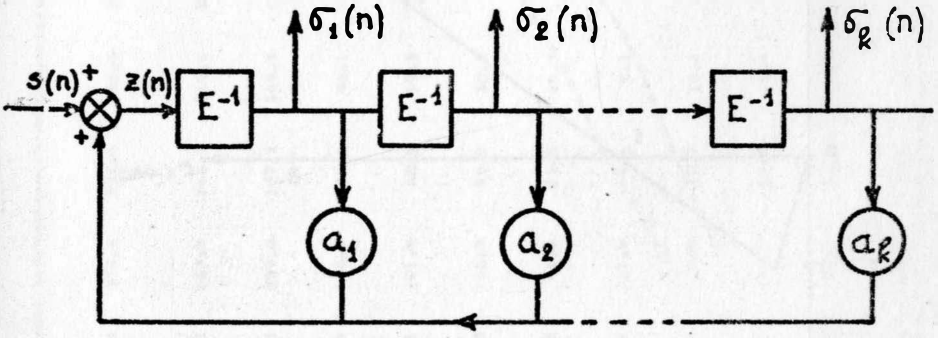
6. Y.TZYPKIN : "Adaptation, Learning and Self Learning in Control Systems" -I.F.A.C. Congress - London 1966.

7. P.EYKHOFF : "Process Parameter Estimation" Progress in Control Engineering - Edited by Mac Millan - Heywood book-London 1964.

8. A.RAULT, R.POULIQUEN, J.RICHALET : "Identification and Sensitivity" - Second I.F.A.C. Symposium on System Sensitivity and Adaptativity - Dubrovnik 1968.



— Fig. 1 Procédure d'identification —



— Fig. 2 —

(E^{-1} opérateur retard)

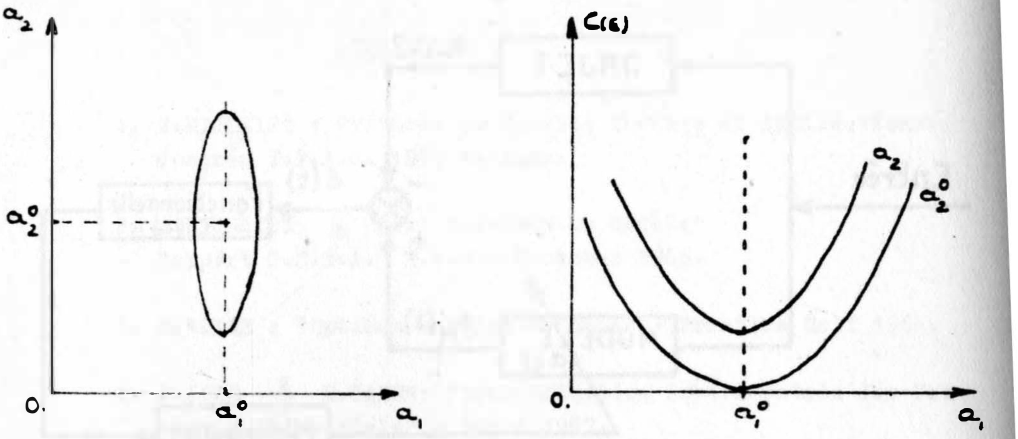


FIG. 3

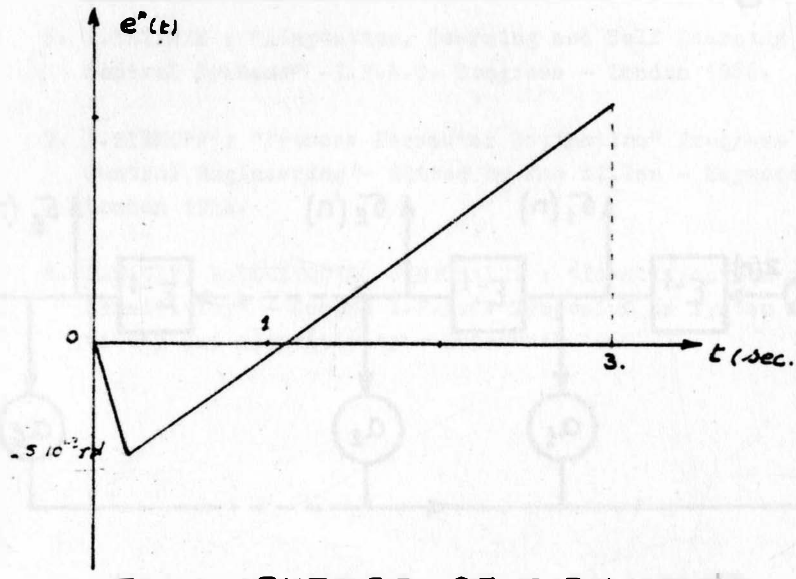


FIG. 4 ENTREE SENSIBILISANTE

A= 0.495

B= C=	12.00	12.50	13.00	13.50	14.00	14.50	15.00	15.50	16.00	16.50	17.00	17.50
-2.50	418.4	378.4	338.4	298.4	258.5	218.4	178.5	138.4	98.5	58.6	18.7	22.9
-2.25	317.0	282.6	248.4	214.1	179.7	145.5	111.1	76.8	42.4	8.3	27.4	61.3
-2.00	236.5	207.0	177.7	148.3	119.1	89.7	60.3	30.9	3.5	28.8	57.9	87.1
-1.75	173.4	148.3	123.1	98.1	73.1	48.1	23.0	4.0	27.9	52.8	77.7	102.7
-1.50	123.1	101.9	80.6	59.3	38.0	16.7	5.4	26.6	47.8	68.9	90.1	111.5
-1.25	84.9	66.9	48.8	30.8	12.8	6.4	24.0	42.0	60.0	77.9	96.0	113.9
-1.00	56.3	41.0	25.8	10.7	7.6	20.5	35.6	50.8	66.0	81.3	96.5	111.7
-0.75	34.7	21.9	9.4	9.3	18.4	29.9	42.6	55.5	68.2	81.0	93.9	106.6
-0.50	19.4	9.3	11.2	18.2	26.8	36.2	46.1	56.5	66.9	77.6	88.4	99.1
-0.25	10.7	13.3	19.0	25.9	33.5	41.4	49.5	57.9	66.4	74.8	83.5	92.2
0.0	15.5	20.6	26.5	32.9	39.6	46.5	53.6	60.7	67.8	75.1	82.5	89.8
0.25	22.4	27.8	33.5	39.4	45.5	51.8	58.0	64.3	70.8	77.3	83.7	90.1

180-10%

Fig 5 SALAM_TANGAGE_ENTREE PILOTEE $\frac{A+Bp}{1+Cp+0.9p^2+1.5p^3}$

A= 0.495

C= \ B=	12.00	12.50	13.00	13.50	14.00	14.50	15.00	15.50	16.00	16.50
-2.00	62.7	66.0	70.8	76.5	82.4	88.7	95.4	102.1	108.9	115.7
-1.75	49.6	51.7	54.9	59.5	64.5	69.9	75.7	81.5	87.5	93.5
-1.50	38.3	38.7	40.5	43.8	47.9	52.5	57.4	62.5	67.6	72.8
-1.25	29.0	27.9	27.7	29.5	32.7	36.3	40.6	44.9	49.3	53.8
-1.00	22.4	19.6	17.6	17.0	18.7	21.6	25.0	28.7	32.5	36.3
-0.75	19.7	15.6	11.8	8.3	6.4	7.9	10.7	13.9	17.6	21.5
-0.50	24.3	21.1	18.2	15.3	12.8	10.8	11.2	13.7	16.7	20.0
-0.25	33.1	30.6	28.2	25.8	23.6	21.9	20.9	21.8	23.5	25.9
0.0	41.4	39.4	37.4	35.4	33.6	32.2	31.2	31.3	32.6	34.3
0.25	48.9	47.4	45.8	44.3	42.8	41.7	40.9	40.8	41.6	43.3

Iso 10%

Fig 6 SALAM - TANGAGE - ENTREE SENSIBILISANTE $\frac{A+Bp}{1+Cp+0.9p^2+1.5p^3}$

PROCESS DYNAMIC IDENTIFICATION BY THE MULTISTEP METHOD

M. Menahem

Societé Contrôle Bailey
32 Bd Henri IV, Paris, France

I. INTRODUCTION

The step-response test,i.e. an experimental deriving of the system step-response is a practical identification method widely used in industrial practice, mainly because of it simplicity and time-economy. In many cases it may stand for the frequency-domain analysis with a pretty good results. Still, the step-response test as well as the frequency analysis may need a lot of research time when applied in its simple form.In practise, one is obliged to repeat the procedure many times to average results, especially when external disturbances are unneglectable. On the other way the repeating is necessary when verifying process linearity hypothesis, by using steps with two or three different magnitudes.

Thus,the step-response test may lead to a considerable loss of time, if the process transient time T_o is large, i.e. in thermal processes. If the experiment´s time interval T is defined and bounded, the unit time T_o determines a total admissible number of trials. But that´s just the main cumber of the method: to keep calculations simple one should´nt repeat trials with the frequency greater than $1/T_o$; conversly,to get more information during defined experiments´time, more trials are necessary.

Last remark defines a goal of the multistep method: to deal with a train of steps, number of which during some defined time-interval is much greater than a possible number of distinot steps, restricted by T_o time.

The multistep method represents generalization of the classical step-response test. The last one should be treated as a simplest case of the method. The step-response test requires to put the process out of action for a larger time, while the test´s sensibility for external disturbances remains considerable. The multistep method enables a becer utilization of the experiments' time i.e. a greater accuracy during defined time-interval.

This paper deals with general principles of the multistep method and its computational aspects. A problem of choice of proper control sequencess to simplify computations will be discusses. A strong relation existing between a computational process in time domain and the discrete Fourier transform will be presented. Last of all, a choice of multiple sequences for multivariable systems identification will be discussed. Actually experiments on boilers in a thermal power station are carrying into effect. The experimental results will be known, when presenting this paper at the Congress.

II. THE MULTISTEP METHOD

Consider a physical system with an input variable e(t) . Let's s(t) be an output variable, which time response due to e(t) is under investigation. We are concerned with the open dynamic system, as well as with the closed-loop control system, which dynamic properties are to be identified. Let's take a train of steps with defined magnitude, applied at determined time-moments as a testing signal.

Suppose T_0 is a unit time, equal to the estimated transient time in the process. Suppose T_0 is known. We can assume that a priori limitation of the unit time is rather a qualitative result issuening from our experience on the process.

T is an active time, i.e. the time interval during which successive steps will be applied.

Δ is an elementary time-interval. Choice of Δ depends on the duration time of the transient processes. Δ determines a control step and an observation step. It determines a number of points describing the seeking process' response during transient time.

Assume that our process response is a step-response, represented in discrete moments of time by $1 + (\Delta/T)$ values. These values, denoted by i_0, i_1, ..., i_m, ($m = T/\Delta$) constitute an index sequence.

The point of the matter is an estimation of the index sequence in some limited time. Furthermore, we assume that our process has a limited static gain, as the identification of processes with pure integration doesn't create any considerable difficulties. We state $i_k = i_m$ for all $k \gg m$. Thus, one should make it sure the chosen unit time covers the total transient interval.

Denote by $e_0, e_1, \ldots, e_N$ ($N = T/\Delta$) a sequence of the input signal levels,while by $a_0,a_1,\ldots,a_N$ an acting sequence, i. e. the sequence of increments applied at the process input during the experiment's time T or so called active time. The each increment can be determined by its amplitude (positive,or negative). Some of them may be equal to zero.

There exists a sequential relation $e_k=e_{k-1}+a_k$.If we choose the initial level as the input signal's reference level, we'll get $e_k = \sum a_j$; $j = 0, 1, \ldots, k$.

Let's denote by $s_0,s_1,\ldots,s_N$ a sequence of the signals observed during the active time. We assume the knowledge of the initial stable state of the observed signal (related to the reference level of the input signal). The values s_k represent deviations from this state.

We'll assume now that a hypothesis on the system's linearity is valid, so the superposition principle can be used. Let us note, that up to date this is the only one hypothesis sufficiently general to practical applications in a broad variety of problems met in industrial practice (except those cases where a considerable nonlinearity does exist).Even if, for certain defined magnitudes' interval some nonlinearities will occur (mainly due to curved process' characteristics), the aim of the identification procedure is the best possible approximation of the real process by a linear model, with sufficient accuracy for the later processs' control.

Let's express the output signal at the moment k as a function of the control signal in current time and in the past

$$s_k = \sum_{j=0}^{m-1} i_j a_{k-j} + e_{k-m} i_m + \mathcal{E}_k \qquad (1)$$

where $\mathcal{E}_k$ denotes an effective error (unknown) at the moment k . $\mathcal{E}_k$ issues from many reasons: disturbances during measurement process, process nonlinearities, measurement inaccuracies, uncorrect estimation of the process transient time.

The relation (1) is just a discrete form (corresponding to the particular signal we're dealing with) of the known equation

$$s(t) = \int_0^{T_0} i(u)\,\dot{e}(t - u)du + e(t - T_0)\,i(T_0) + \mathcal{E}(t). \quad (2)$$

here $\dot{e}(t)$ - time derivation of $e(t)$. One should keep in mind an assumption on the process finite memory T_0.

In each time moment we may compute the index sequence, taking into account m equations of the form (1), corresponding to moments $k, k-1, \ldots, k-m+1$.

We shall call it the local computation with index k. Each local computation will be performed using assumption on errors $\mathcal{E}_k$ equal to zero. Thus one can easy eliminate this variable from eq. (1).

A set of these linear equations can be written in vector form

$$s(k) = H(k - 2m + 2)i + e(k - m)i_m \quad (3)$$

with following notations

$s(k) = (s_{k-m+1}, \ldots, s_{k-1}, s_k)'$ - the observation vector at the moment k

$i = (i_{m-1}, \ldots, i_1, i_0)'$ - the index vector

$e(k - m) = (e_{k-2m+1}, \ldots, e_{k-m-1}, e_{k-m})'$

$$H(k - 2m + 2) = \begin{bmatrix} a_{k-2m+2} & a_{k-2m+3} & \cdots & a_{k-m+1} \\ a_{k-2m+3} & a_{k-2m+4} & \cdots & a_{k-m+2} \\ \vdots & \vdots & & \vdots \\ a_{k-m+1} & a_{k-m+2} & \cdots & a_k \end{bmatrix}$$

$s(k)$, i, $e(k - m)$ are column vectors with m components, $H(k - 2m + 2)$ - square matrix of dimension m, with well known structure. Such a matrix is called Hankel matrix [1] .One can note it is symmetric so diagonalization is always possible.

If the inverse of the matrix $H(k - 2m + 2)$ exists,the local computation with index k can be performed. Considering the measurement interval of lenght m one can derive the index sequence, using equation

$$i = H^{-1}(k - 2m + 2)(s(k) - e(k - m)i_m) \tag{4}$$

We stress on two important properties:

1) From the first sight-seeing it lookes like the number of relations in the vector equation (3) is unsufficient to compute the full index sequence. However, one can recognize that if the number of values admitted for computation of the sequence is sufficient,. i_m will slightly differ from i_{m-1}. If the solution (3) is possible, the result (eq. (4)) can be presented by m relations of the form

$$i_j = b_j - c_j i_m, \quad j = 0, 1, \ldots, m-1 \tag{5}$$

From the last one we may compute i_m , substituting $i_m = i_{m-1}$. Such a hypothesis is admissible if the length of the admitted unit time is large enough. Besides, one can recognize that from the relation (5) for $j = 0$ yields a different method for computing i_m , true for most of the real systems.

2) It is evident that each local computation depends only on the Hankel matrix inverse. If, for a defined measurement interval the inverse of the corresponding Hankel matrix doesn´t exist, no information can be derived from that interval.

On the other hand, the each matrix H directly corresponds to the applied control sequence. So, the information, one can obtain from the measurement period by means of many successive local computations, depends on properties of the matrix H, issued from the sequence a_j .

So the problem is not to content oneself with a choice of arbitrary control sequence, but to improve a fact we have a possibility of a free choice of the magnitudes a_j , at least in a broad interval to obtain special sequences,such that corresponding Hankel matrices are invertible.

We emphasize that obtaining of the local results depends only on the control sequence, not on the output measurements quality. It's evident however, that direct influence of the measurements quality as well as of the disturbance level on the results dispersion does exist.

A representation of the sequence a_j presents Fig. 1. It explains a type of Hankel matrix issuing from the sequence. As one may recognize, the sequence can be expressed as a sequence of the column vectors $x(0), x(1), \dots, x(N-m+1)$ with m components $a_k, a_{k+1}, \dots, a_{k+m-1}$. Each vector $x(k)$ includes only one new scalar variable, when compared with the previous vector $x(k - 1)$.

The choice of the sequence a_j rely on the following remark: each of the vectors $x(k)$ can be obtained by a linear transformation on the previous one. The transformation is described by a square matrix A of dimension m, i.e.

$$x(k) = A\, x(k - 1) \; ; \quad A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ & & \ddots & & \\ & & & & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_m \end{bmatrix} \tag{6}$$

This operator is regular when $\alpha_1 \neq 0$.

On the other hand, each Hankel matrix resulting from the sequence a_j consists of m successive vectors. Linear independence of the vectors guaranties existence of the inverse of the corresponding matrix H. One can always choose the only one operator A such that a sequence of the m vectors

$$x, Ax, A^2x, \dots, A^{m-1}x \tag{7}$$

constitutes a set of linear-independent vectors.

So, the control sequence should be determined in virtue of the following:

1) An initial vector $x(0) = (a_0, a_1, \dots, a_{m-1})'$, called a basic sequence which determines first m increments of the control sequence.

2) A regular operator A in canonical form stated from (6).

The choice of the basic sequence will be made taking into account some restrictions, e.g.

- restrains on the steps magnitude,
- trend to enrich an identification signal in the frequency interval corresponding to anticipated use of the system´s dynamic characteristics,
- basic fact, that A and x(0) cannot be chosen independently one after another, when demand for systems of type (7) to be free.

If we assume for matrix A be diagonalizable, x(0) must be an integer linear combination of eigenvalues of the matrix. A . Denoting them by $v_1, v_2, \ldots, v_m$, a condition for x(0) will be

$$x(0) = z_1 v_1 + z_2 v_2 + \ldots + z_m v_m \quad \text{where} \quad z_i \neq 0 \quad \text{for all} \quad i$$

It´s easy to recognize a considerable simplification of the computation issuing from such a control sequence: nevertheless, from the considered measurement interval (of lenght m), the matrix inverse always concerns the same Hankel matrix, independently of a power of the operator A (let´s note that the matrix A inversion doesn´t create any troubles). This only matrix is

$$H(0) = \left[x(0), Ax(0), \ldots, A^{m-1}x(0) \right] \tag{8}$$

Just a first matrix to be inverted is H(0). It consists of m observations

$$s_{m-1}, s_m, \ldots, s_{2m-2}$$

the last are to be inverted is a matrix

$$H(N - 2m + 2) = A^{N-2m+2} H(0) \tag{9}$$

where N denotes the control sequence length.

This matrix corresponds to observations

$$s_{N-m+1}, s_{N-m+2}, \ldots, s_N$$

A total number of the index sequences obtained from the computation, demanding to invert H(0) only, equals $N - 2m + 3$. When $N = nm$, the number of these sequences equals $(n - 2)m + 3$; while for the same observation period the classical step-response method allows us to apply only n successive steps, distant for a time T_0 one from another, i.e. n repetitions only are possible. For example, for $m = 20$, $n = 4$ the multistep method equivalents to 43 steps while the classical method would allow for 4 only.

It may be proved that inversion of the matrix of type H(0) doesn't creat any troubles. More precisely, the inversion associates with the inversion of Vandermonde's matrix. In fact, consider a square matrix of dimension m

$$H = \left[x,\ Ax,\ A^2x,\ \dots,\ A^{m-1}x\right] \qquad (10)$$

where the vector v and the operator A are given. If we assume that A is regular, of the canonical form of (6) and possesses distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_m$, then it can be written in a known form

$$A = VDV^{-1} \qquad (11)$$

where $D = \left\{\lambda_1, \lambda_2, \dots, \lambda_m\right\}$ - diagonal matrix

$$V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_m \\ \vdots & \vdots & & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \dots & \lambda_m^{m-1} \end{bmatrix} \text{ - Vandermonde matrix}$$

It yields

$$H = \left[x,\ VDV^{-1}x,\ VD^2V^{-1}x,\ \dots,\ VD^{m-1}V^{-1}x\right] = \\ = V\left[y,\ Dy,\ D^2y,\ \dots,\ D^{m-1}y\right] \qquad (12)$$

where $y = V^{-1}x = (y_1, y_2, \dots, y_{m-1})'$ (13)

On the other hand, stating

$$D^k y = Y_{d_k} \qquad (14)$$

where $Y = \{y_0, y_1, \dots, y_{m-1}\}$ - diagonal matrix and $d_k = (\lambda_1k, \lambda_2k, \dots, \lambda_mk)'$ equation (12) will be

$$h = VY\left[d_0, d_1, d_2, \dots, d_{m-1}\right] = VYV' \qquad (15)$$

Thus, the matrix H has been partitioned to a product of matrices with known inverses. Let us note, the regularity of matrix H demands for the vector $V^{-1}x$ to have no components equaled to zero.

III. PARTICULAR, FUNDAMENTAL CASE

It seems like the simplest procedure for obtaining the control sequence rely upon a choice of the matrix A in the form of

$$A = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ & & & \ddots & 1 \\ -1 & 0 & 0 & \dots & 0 \end{bmatrix} \qquad (16)$$

We assume, for all of the future consideration that m <u>is even</u>. Then det (A) = +1. One can check that A is an orthogonal operator, so $AA' = I$.

Such a choice of the matrix A leads up to periodic control sequence with a period equal to the double unit time (see Fig. 2).

If a vector

$x(0) = (1, 0, \dots, 0)'$

is applied as the basic sequence, the problem reduces to the classical step-response method (a train of steps with equal magnitudes, while signs changing alternatively).

For arbitrary basic sequence the matrix H(0) is of a known form, easy to compute inverse. Before starting the computation, let's transform H(0) to the more familiar form, by a simple column displation

$$C(0) = H(0)\cdot Z \quad \text{where} \quad Z = \begin{bmatrix} 0 & \cdots & 1 \\ & \diagup & \\ 1 & \cdots & 0 \end{bmatrix} \tag{17}$$

Hence

$$C(0) = \left[A^{m-1}\,x(0),\ \ldots,\ A\,x(0),\ x(0) \right] \tag{18}$$

$$C(0) = \begin{bmatrix} a_{m-1} & a_{m-2} & \cdots & a_1 & a_0 \\ -a_0 & a_{m-1} & \cdots & a_2 & a_1 \\ -a_1 & -a_0 & \cdots & a_3 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ -a_{m-2} & -a_{m-3} & \cdots & -a_0 & a_{m-1} \end{bmatrix} \tag{19}$$

The computation of the eigenvalues and eigenvectors of C(0) is evident. Thus only final results will be presented.

First, note that the characteristic equation for A is

$$\lambda^m + 1 = 0 \tag{20}$$

Denote different roots of the Eq. (20) by $\lambda_1, \lambda_2, \ldots, \lambda_m$. The m of the eigenvalues of $C(0)$ will be expressed by the relation

$$y_k = a_{m-1} + \lambda_k a_{m-2} + \lambda_k^2 a_{m-3} + \cdots + \lambda_k^{m-1} a_0 \tag{21}$$

$$\left(k = 1, 2, \ldots, m \right)$$

To each eigenvalue corresponds the eigenvector

$$v_k = (1,\ \lambda_k,\ \lambda_k^2,\ \ldots,\ \lambda_k^{m-1})' \tag{22}$$

It may be checked, that $y_k v_k = C(0)\cdot v_k$

Denote

$$V = \left[v_1,\ v_2,\ \ldots,\ v_m \right] \tag{23}$$

where V represents a matrix composed from the eigenvectors of C(0) . It results then

$$C(0) = VYV^{-1} \text{ , where } Y = \left\{y_1, y_2, \ldots, y_m\right\} \tag{24}$$

C(0) can possess invers if and only if none of the eigenvalues y_k equals zero. This regularity condition defines at the same time <u>the validity condition for the basic sequence</u> x(0). The last condition can be formulated as follows:

Let's $x(0) = (a_0, a_1, \ldots, a_{m-1})'$ be a basic sequence. The necessary and sufficient condition for utilization of the sequence is that the set of algebraic equations:

$$\begin{cases} Z^m + 1 = 0 \\ a_0Z^{m-1} + a_1Z^{m-2} + \ldots + a_{m-1} = 0 \end{cases} \tag{25}$$

do not possess any common roots.

If the condition (25) is satisfied, $C^{-1}(0)$ exists and equals

$$C_0^{-1} = VY^{-1}V^{-1} \tag{26}$$

Note that the inverse Vandermonde matrix can be obtained easy. Thus a relation for V^{-1} is particulary simple because of particular values of the components of V . In fact , it's easy to check that

$$V^{-1} = \frac{1}{m} ZV' \tag{27}$$

where V' is transponese of V , Z is defined by Eq. (17).

Note that the eigenvalues of A are expressed by

$$\lambda_{k+1} = \exp\left(j(2k+1)\,\pi/m\right) \text{ , } k = 0, 1, \ldots, m-1 \tag{28}$$

and

$$\lambda_{k+1} = \bar{\lambda}_{m-k} \tag{29}$$

(where $\bar{\lambda}$ denotes a complexe number, conjugate with λ).

From Eq. (21) it also results

$$y_{k+1} = \bar{y}_{m-k}, \quad k = 0, 1, \ldots, m-1 \tag{30}$$

On the other hand it's easy to show that the invers of C(0) is a matrix of the same type than C(0). This property can be expressed by

$$\left[A^{m-1}x(0), \ldots, Ax(0), x(0)\right]^{-1} =$$

$$= \left[A^{m-1}u(0), \ldots, Au(0), u(0)\right] \qquad (31)$$

The components of the vector u(0) existing in the inverse of C(0) can be expressed by components of x(0). Stating

$$u(0) = (u_0, u_1, \ldots, u_{m-1})' \qquad (32)$$

the component u_k is given by

$$u_k = \frac{2}{m}\sum_{h=1}^{m/2}\frac{\mathrm{Re}\,(y_h\lambda_h^{m-k-1})}{y_h\,\bar{y}_h} \qquad (33)$$

where y_h is given by Eq. (21), Re (a) = real part of a .

Thus the inverse of the matrix C(0) is expressed in a explicit form. Therefore the direct inversion by a computer and, consequently, considerable computation errors (especially when dimension m is large) can be avoided.

To complete our considerations we'll express the index vector in a direct form. To do this, consider once more the matrix H(0). In fact, as $Z^2 = I$, the Eq. (3) can result as fol lowing

$$s(k) = H(k - 2m + 2)Z^2 i + e(k - m)i_m =$$

$$= H(k - 2m + 2)Z\,\hat{i} + e(k - m)i_m =$$

$$= C(k - 2m + 2)\,\hat{i} + e(k - m)i_m \qquad (34)$$

where

$$\hat{i} = Zi = (i_0, i_1, \ldots, i_{m-1})' \qquad (35)$$

and

$$C(k - 2m + 2) = A^{k-2m+2}\,C(0) \qquad (36)$$

Hence we have

$$C^{-1}(k - 2m + 2) = C^{-1}(0)\,(A)^{k-2m+2} \quad \text{as } \Lambda\Lambda' = I \qquad (37)$$

Because of

$$C_0^{-1} = \begin{bmatrix} \hat{u}'(0) \\ \hat{u}'(0)\, A \\ \vdots \\ \hat{u}'(0)\, A^{m-1} \end{bmatrix} \tag{38}$$

where

$$\hat{u}(0) = Zu(0) = (u_{m-1}, \ldots, u_1, u_0)' \tag{39}$$

we obtain m scalar relations of the form

$$i_l = \hat{u}'(0)\, (A')^{k-2m+2+l} (s(k) - e(k-m)i_m) \tag{40}$$

$$l = 0, 1, 2, \ldots, m-1$$

where i_l is an l-nt component of the index vector. Remind that Eq. (40) for $l = 0$ and $l = m - 1$ permit to obtain i_m . Actually the local computation with an index number k (i. e. the computation of the second component of the Eq. (40)) never results i_l but

$$i_l + \tilde{\mathcal{E}}_{lk}$$

where $\tilde{\mathcal{E}}_{lk}$ represents an l-th component of the vector

$$C^{-1}(k - 2m + 2) \cdot \mathcal{E}(k) \tag{41}$$

$\mathcal{E}(k) = (\mathcal{E}_{k-m+1}, \ldots, \mathcal{E}_{k-1}, \mathcal{E}_k)'$ - a vector with components representing observation errors made in the considered measurement interval (see Eq. (1)).

<u>REMARK</u>. It can be shown that the particular easiness of the presented computation scheme results from a relationship existing between the proposed method and <u>the discrete Fourier transform</u>.

The method can be applied to identification of a discrete "impulse" string describing a dynamic system.Noting that,let's come back to Eq. (1)

$$s_k = i_0 a_k + i_1 a_{k-1} + \ldots + i_{m-1} a_{k-m+1} + i_m e_{k-m} + \mathcal{E}_k$$

Applying a recurrence formula $e_k = e_{k-1} + a_k$ it's possible to evaluate s_k by the expression

$$s_k = d_0 e_k + d_1 e_{k-1} + \dots + d_m e_{k-m} + \mathcal{E}_k \tag{42}$$

where

$$d_j = i_j - i_{j-1} \quad (d_0 = i_0) \tag{43}$$

In particular, the sequence $d_0, d_1, \dots, d_{m-1}$ is called a discrete impulse string. The sequence i_j exactly represents a discrete impulse response $i(t)$. However, it doesn't represent exactly the continous impulse response. The sequence d_j can be treated rather as a modified representation of the discrete index sequence.

Now, in place of the Eq. (1), consider a new equation

$$s_k = \sum_{0}^{m-1} d_j e_{k-j} + \mathcal{E}_k \tag{44}$$

assuming d_j equals zero for all $j \geqslant m$.

The Eq. (34) admits the form

$$s(k) = C_e(k - 2m + 2)d + \mathcal{E}(k) \tag{45}$$

where $d = (d_0, d_1, \dots, d_{m-1})'$ (46)

The index e reminds of the fact that C comprises now two input levels e_j instead of increment magnitudes a_j. Hence

$$C_e(0) = \left[A^{m-1}\, e(0), \dots, A\, e(0), e(0) \right] \tag{47}$$

where

$$e(0) = (e_0, e_1, \dots, e_{m-1})' \tag{48}$$

The computation procedure aiming to compute the vector d certainly remains the same as before. The application of these new denotations allows, however, for a better presentation of a parallels existing between the procedure and utilizatization of the discrete Fourier transform.

Let's state a time sequence with finited length

$$x_0, x_1, \ldots, x_{m-1} \quad (m \text{ is event })$$

Mostly, the discrete Fourier transform of this sequence forms a new sequence [3]

$$X_0, X_1, \ldots, X_{m-1}$$

where

$$X_r = \sum_{0}^{m-1} x_k (W^r)^k \qquad (49)$$

$$W = \exp(-j \cdot 2\pi/m), \; r = 0, 1, \ldots, m-1$$

Consider first the sequence

$$X_r = \sum_{0}^{m-1} x_k (W^r + 1/2)^k \qquad (50)$$

It represents a connection between the transformation of the complex sequence $x_k W^{k/2}$ (in sense of the def. (49)) and the real sequence x_k.

Utilizing denotations (28) we'll get

$$W^{r+1/2} = \lambda_{r+1}^{-1} = \lambda_{m-r} \qquad (51)$$

Thus we can write over

$$X_r = \sum_{0}^{m-1} x_k \lambda_{m-r}^{k} \qquad (52)$$

or in vector notation

$$X = \begin{bmatrix} 1 & \lambda_m & \lambda_m^2 & \cdots & \lambda_m^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{m-1} \end{bmatrix} \cdot x \qquad (53)$$

where $X = (X_0, X_1, \ldots, X_{m-1})'$ and $x = (x_0, x_1, \ldots, x_{m-1})'$.

Application of the def. (23) and Eq. (27) results in the relationships

$$X = mV^{-1}x \qquad (54)$$

or

$$x = \frac{1}{m} VX \qquad (55)$$

The relation (55) states an analogy of the inverse Fourier transform.

Consider now an operation of the "finite" convolution of the form appearing in Eq. (44). It can be represented by the vector equation

$$s = Cd \tag{56}$$

where C is of the form (47).

Transforming both sides of the (56) yields

$$mV^{-1}s = mV^{-1}Cd = V^{-1}CV \cdot mV^{-1}d \tag{57}$$

or denoting $D = mV^{-1}d$ and $S = mV^{-1}s$ (58)

$$S = V^{-1}CVD \tag{59}$$

But, according to Eq. (24), $V^{-1}CV = Y$ and to Eq. (21), an each element of the diagonal matrix Y can be presented in the form

$$y_j = e_0\lambda_{j+1}^{m-1} + e_1\lambda_{j+1}^{m-2} + \dots + e_{m-1}, \quad j = 0,1,2,\dots,m-1 \tag{60}$$

At last, denoting the transformation of the basic sequence by

$$E = mV^{-1}\, e(0) \tag{61}$$

it´s easy to check

$$y_j = \lambda_{j+1}^{m-1} E_j \tag{62}$$

Hence, the vector equation (59) can be written in a form of the m independent scalar equations

$$\lambda_{j+1}^{m-1} S_j = E_j D_j \,, \quad j = 0,\ 1,\ \dots,\ m-1 \tag{63}$$

Considering Eq. (63) one should remember,that in the transformation (54) convolution operation is substituted by a usual product, while S_j indicates for a time shift between output and input sequences (note that periodicity of the output begins no more then in the (n - 1)th moment).

This property clarifies the computation easiness we have obtained by a choice of the particular operator A ,in the first part of this paragraph. It´s of a significant importance however, that the gathered results are obtained on the base of time-domain considerations only. Still, there exists a natural relationship with the classical deconvolution method based on the mathematically justified application of the transformation considered.

IV. CHOICE OF THE MULTIPLE SEQUENCES

The multistep method can be, without any considerable difficulties generalized for multiinput systems. Our aim, as before is to develop a method enableing a better utilization of the time provided for research. Increasing the frequency of the control signals applied to a given input, passive intervals are narrowed. Hence, the awaiting for ending of the research, regarding these signals responses can be avoided, before starting the new research concerning signals applied to other inputs. Just such the awaiting is unavoided when using the step method. We assume here, althought it´s not enough that the computation easiness will serve as a criterion for the control sequences choice.

Consider as an example, a system with two inputs $e^1(t),e^2(t)$ affecting on the output $s(t)$. Denote by d_j^1 and d_j^2 impulse sequences (in the sense of the condition (45)), describing the system. We make an assumption on the equality of the unit time for the both transient states considered, i.e. $m_1 = m_2$.

The Eq. (44) will be

$$s_k = \sum_{0}^{m-1} d_j^1 e_{k-j}^1 + \sum_{0}^{m-1} d_j^2 e_{k-j}^2 + \varepsilon_k \tag{64}$$

The relation (64) differs from Eq. (44) only by a greater number of parameters to be identified. Denote by

$$d = \left[(d^2)', (d^1)' \right]'$$

the vector of dimension 2m . Utilizing for a local computation with an index k , 2m observations corresponding to the moments

k, k-1, k-2, ..., k-2m+1

we'll obtain a vector equation of type (45):

$$s(k) = C_e(k - 4m + 2)d + \mathcal{E}(k) \qquad (65)$$

where s(k), $\mathcal{E}(k)$ are vectors of dimension 2m, while C_e is a square matrix of dimension 2m. The matrix C_e will be defined however, when considering a relation existing between the sequences

$$e_0^1, e_1^1, \ldots, e_N^1 \text{ and } e_0^2, e_1^2, \ldots, e_N^2$$

It's easy to check that in order to give for the matrix C_e the form derived in item III, the relation should result in

$$e_{k-m}^2 = e_k^1 \text{ and/or } e_k^2 = e_{k+m}^1 \qquad (66)$$

Thus we have two identical sequences shifted in time for a number of intervals equaled m (the lead of e^2 with regard to e^1 issues only from the particular choice of the vector d). Thus the application of the method depends on the initial choice of the basic sequence e(0) of length 2m , with components

$$e_0, e_1, \ldots, e_{2m-1}$$

The sequence e_j^1 expressed in a vector form (as indicated in paragraph II) will be represented by the vectors

$$e(0), A\, e(0), A^2\, e(0), \ldots$$

while for the sequence e_j^2

$$A^m e(0), A\left[A^m e(0)\right], A^2\left[A^m e(0)\right], \ldots$$

where A is an operator of the form (16) with dimension 2m.

Evidently in these conditions we find the same computation reductions we have found in the case of the one-input system.We emphasize the main goal of the choice we made, was to simplify computations. However, we still have a greater number of the degrees of freedom, as the only constraint given on the basic sequence is the constraint of the form (25).

V. CONCLUSIONS

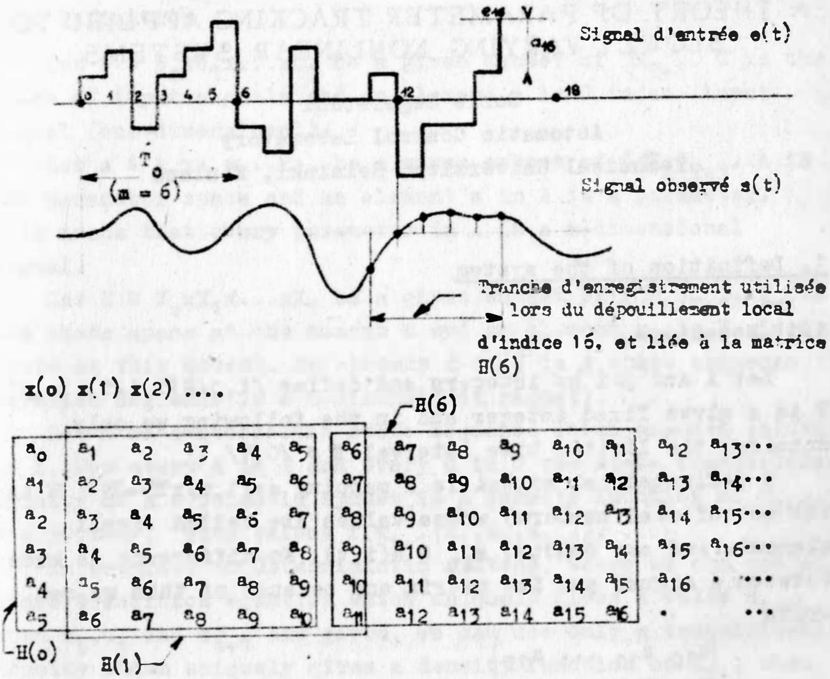
The paper presents an identification method called "the multistep method". It generalizes the classical step-response method. The main advantage of the multistep method is a better utilization of the bounded time designated for identification research. An identification signal used is not an ordinary step but a sequence of steps with defined magnitude,applied at stated intervals. It has been showed, that a computation of the observed signal (to get a required information on the transient properties) can be considerably simplified if a proper choice of the control signals with stated properties is made. The computation doesn´t need a direct inversion of the matrix. Hence results the computation reduction as well as significant errors are avoided (particulary when dealing with matrices of a large dimension: from 10 to 30 for one input variable). Besides, the proper sequence choice enables to accomplish computation based on the observed signal measurements,while eliminating all singularities.Thus all conditions for a large number of the local computations are ensured.The number is justified mainly for a frequent affecting of the control signal on the system input.

There was no place to consider the problem of the statistical estimation of the index or impulse sequences, based on the successive local computations. We only note that the computation results can be of a greater value and of a more general sense than a direct application of the global least squares method to the total measurement accomplished during a research time

* This work is a part of a broader research program under patronage and support of the Fonds de la Recherche de la Délégation Générale à la Recherche Scientifique et Technique ("Automatisation").

REFERENCES

1. F.R. Gantmacher, The Theory cf Matrices. Chelsea, New York (1959), vol. 1 p. 338.
2. R. Bellman, Introduction to Matrix Analysis. Mc Graw Hill (1960) p. 234.
3. W.T. Cochran et al., What is the Fast Fourier Transform? . Proc. IEEE vol. 55 Oct. 1967, p. 1664.
4. M. Menahem, Méthodes Expérimentales,Anciennes et Nouvelles, d'Analyse Dynamique des Systèmes Industriels. Papers of the Congress MESUCORA 1967 (Paris , 17 April 1967).



$$\begin{array}{cccccccccccccc}
a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13}\cdots \\
a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14}\cdots \\
a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15}\cdots \\
a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16}\cdots \\
a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \cdot & \cdots \\
a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & \cdot & \cdots
\end{array}$$

Fig. 1. Vector representation of the acting sequence
Input signal e(t)
Observed signal s(t)
Measurement interval utilized in the local computation with index 16, corresponding to matrix H(6)

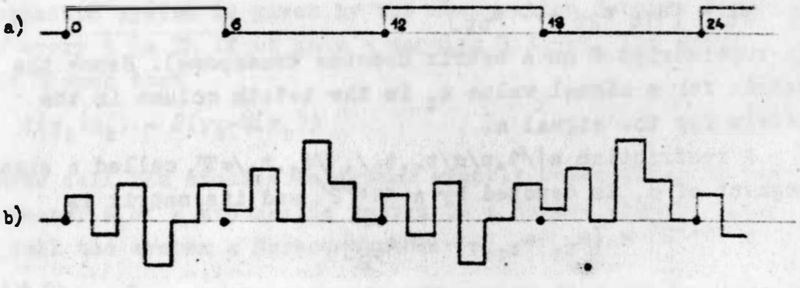


Fig. 2. Periodic sequence (period = 2 m)
basic sequence: a) (1, 0, 0, 0, 0, 0)
b) (1, -3/2, +5/2, -3, +3/2, +3/2)

A THEORY OF PARAMETER TRACKING APPLIED TO SLOWLY VARYING NONLINEAR SYSTEMS

Boris Segerståhl
Automatic Control Laboratory
Technical University, Helsinki, Finland

1. Definition of the system

1.1. Notations

Let i and $j>i$ be integers and define $/i,j/ \triangleq \{i,i+1,\dots,j\}$. T is a given fixed integer and in the following we only concider the limited time interval $\mathbb{T} \triangleq /0,T/$.

A n-dimensional signal is a mapping $a:/1,n/ \times \mathbb{T} \rightarrow \mathbb{R}$ ($\mathbb{R}$ is the set of real numbers) whose values are called signal elements and are denoted $a_{it} \triangleq a(i,t)$. No difference is made between a signal and its matrix and because of this we can write

$$a = \begin{bmatrix} a_{10} & a_{11} & \cdots & a_{1T} \\ a_{20} & a_{21} & \cdots & a_{2T} \\ \vdots & \vdots & & \vdots \\ a_{n0} & a_{n1} & \cdots & a_{nT} \end{bmatrix}$$

The signal value a_t associated with a signal a is a mapping $a_t:/1,n/ \rightarrow \mathbb{R}$ such that its matrix is

$$a_t = [a_{1t} \; a_{2t} \; \cdots \; a_{nt}]^T$$

(a superscript T on a matrix denotes transpose). Hence the matrix for a signal value a_t is the t+1:th column in the matrix for the signal a.

A restriction $a|/1,n/ \times /t_1,t_2/$, $/t_1,t_2/ \subset \mathbb{T}$, called a signal segment of a, is denoted by $a^{/t_1,t_2/}$ and its matrix is

$$a^{/t_1,t_2/} = [a_{t_1} \vdots a_{t_1+1} \vdots \cdots \vdots a_{t_2}].$$

In cases where $t_1=0$ we use the shorter notation $a^t = a^{/0,t/}$.

A tilde distinguishes a random variable from its values. Hence $\tilde{x}$ can be a random variable and x one of its values.

1.2. Definition

Let $U \triangleq U_0 x U_1 x \ldots x U_T$ be a given subset of $\mathbb{R}_{T+1}$. U is the space of input signals and an element u in U is an input signal (one-dimensional).

Let $A \triangleq A_0 x A_1 x \ldots x A_T$ be a given subset of $(\mathbb{R}_m)_{T+1}$. A is the parameter space and an element a in A is a parameter. This means that every parameter in A is a m-dimensional signal.

Let $X \triangleq X_0 x X_1 x \ldots x X_T$ be a given subset of $(\mathbb{R}_n)_{T+1}$. X_t is the state space at the moment t and an element x_t in X_t is a state at this moment. An element x in X is a state sequence (a state sequence is a n-dimensional signal).

Let $\tilde{x}$ be a stochastic state sequence, which assumes values in X. For every a in A and every u in U the state transitional density of a stochastic system is a density function of x_{t+1}, $t \in /0,T-1/$, with values $f(x_{t+1}|x_t,u_t,a_{t+1})$.

In contrast to deterministic systems, where we can use a state transition equation which uniquely gives a value x_{t+1} when x_t, u_t and a_{t+1} are given, we can use only a transitional density which uniquely gives a density function on X_{t+1} when x_t, u_t and a_{t+1} are given.

Let $Y \triangleq Y_0 x Y_1 x \ldots x Y_T$ be a given subset of $\mathbb{R}_{T+1}$. Y is the space of output signals and an element y in Y is an output signal (one-dimensional).

The relation between state and output signal value in a stochastic system is given by the observation density $f(y_t|x_t)$ for every t in $\mathbb{T}$. If we know a mapping $G:X_t \rightarrow Y_t$ for every t in $\mathbb{T}$ such that

$$f(y_t|x_t) = \delta(y_t - G(x_t)) \qquad (1)$$

thenwe call the stochastic system exactly observable. If the parameter a of a stochastic system is a Markov signal, then we call the system a Markov system.

These notations give the following characterization of a Markov system on the time interval $\mathbb{T}$.

A Markov system on the interval T is defined by:

1) The initial state density $\hat{f}_{-1}(x_0|a_0)$ and the initial parameter density $\hat{f}_{-1}(a_0)$,
2) The state transitional density $f(x_{t+1}|x_t,u_t,a_{t+1})$ and the parameter transitional density $f(a_{t+1}|a_t)$,
3) The observation density $f(y_t|x_t)$.

In the characterization given above the initial conditions usualy given for deterministic systems have been replaced by initial densities.

The general difference between deterministic and stochastic systems is hence that when we use values of functions in deterministic systems we use density functions in stochastic systems and these density functions are as unique as the values of the functions in the deterministic case.

2. Formulation and solution of the problem

2.1. Formulation

Our problem is a tracking problem and hence our main interest is concentrated on the density function of the Markov parameter at every moment $t = 0,1,\ldots,T$. This density is at every moment conditional with respect to an input signal segment u^{t-1} and an output signal segment y^t. Formaly the problem can be formulated:

> Let a Markov system be defined by its state transitional density, parameter transitional density and observation density. Let the initial state density and initial parameter density be given. Let u be a given input signal and y a given output signal.
> Determine at every moment $t \in /0,T-1/$ the conditional density function
>
> $$f(a_{t+1}|u^t,y^{t+1}) \hat{=} f'_{t+1}(a_{t+1}). \qquad (2)$$

2.2. Solution

The problem formulated on the preceding page is a typical bayesian problem and the solution can be obtained in the following way:

The following density functions are given,

1) state transitional density $f(x_{t+1}|x_t,u_t,a_{t+1})$,
2) initial state density $\hat{f}_{-1}(x_o|a_o)$,
3) parameter transitional density $f(a_{t+1}|a_t)$,
4) initial parameter density $\hat{f}_{-1}(a_o)$,
5) observation density $f(y_t|x_t)$.

From 2) and 4) can be computed the initial joint density

$$\hat{f}_{-1}(x_o,a_o) = \hat{f}_{-1}(x_o|a_o)\hat{f}_{-1}(a_o). \tag{3}$$

Let the first observed output signal value be y_o. The joint conditional density of x_o,a_o is

$$f_o(x_o,a_o|y_o) \triangleq f'_o(x_o,a_o) = \hat{f}_{-1}(x_o,a_o)f(y_o|x_o)N_o(y_o) \tag{4}$$

where $N_o(y_o)$ is the normalizing constant (for given y_o) defined by the condition

$$N_o(y_o) = \left[\int_{X_oxA_o} \hat{f}_{-1}(x_o,a_o)f(y_o|x_o)dx_oda_o\right]^{-1}. \tag{5}$$

From (4) one can immediately compute $f'_o(a_o)$ as

$$f'_o(a_o) = \int_{X_o} f'_o(x_o,a_o)dx_o. \tag{6}$$

The following step is to predict the density function of x_1,a_1 for given u_o. This predicted density function is given by

$$f_o(x_1,a_1) = \int_{X_oxA_o} f(x_1|x_o,u_o,a_1)f(a_1|a_o)f'_o(x_o,a_o)dx_oda_o. \tag{7}$$

After this the procedure outlined above can be continued recursively for every t in $\mathbb{T}$ and the general algorithm at a moment t is:

1) Compute

$$\hat{f}_t(x_{t+1},a_{t+1})=\int_{X_t x A_t} f(x_{t+1}|x_t,u_t,a_{t+1})f(a_{t+1}|a_t)f_t'(x_t,a_t)dx_t da_t$$

2) measure y_{t+1} and compute

$$f_{t+1}'(x_{t+1},a_{t+1}) = \hat{f}_t(x_{t+1},a_{t+1})f(y_{t+1}|x_{t+1})N_{t+1}(u^t,y^{t+1})$$

where

$$N_{t+1}(u^t,y^{t+1}) = \left[\int_{X_{t+1} x A_{t+1}} f_t(x_{t+1},a_{t+1})f(y_{t+1}|x_{t+1})dx_{t+1}da_{t+1}\right]^{-1}$$

3) compute

$$f_{t+1}'(a_{t+1}) = \int_{X_{t+1}} f_{t+1}'(x_{t+1},a_{t+1})dx_{t+1}$$

This general bayesian algorithm is theoreticaly almost trivial and analogous to corresponding metods for state estimation[1], but computationaly it is in general not easy to realize.

Even in the linear case with Gaussian densities it is difficult to use this algorithm for other systems than exactly observable ones where we can find an a posteriori state mapping $H:X_{t-1}xU_{t-1}xY_t \longrightarrow X_t$ which gives the rule for computing the value of the state vector at every moment t.

The main difficulty is due to the fact that the output signal values even in a linear system will be products of parameter values and state values and if we are given only density functions for parameter and state values we have to compute density functions of products of stochastic variables and this is rarely an easy task.

3. The solution for exactly observable linear systems

The necessary density functions and mappings for exactly observable linear Gauss-Markov systems can be constructed in the following way.

Let m and n be positive integers defining the order of the system in such a way that the a posteriori state mapping can be constructed by using the rule that for every t in $\mathbb{T}$

$$x_t = \left[u_t \; u_{t-1} \cdots u_{t-m} \; y_t \; y_{t-1} \cdots y_{t-n}\right]^T \tag{8}$$

where u_i and y_i are input and output signal values. All signal values with negative time index are initial values and can be included in an initial state vector x_{-1}.

Because u_t is included in x_t and because we can construct x_{t+1} if we know u_{t+1}, y_{t+1} and x_t we can replace the state transitional density $f(x_{t+1}|x_t,u_t,a_{t+1})$ by the equaly informative density $f(y_{t+1}|x_t,a_{t+1})$, and we assume that

$$f(y_{t+1}|x_t,a_{t+1}) = c\exp\left\{-\mu(y_{t+1} - x_t^Ta_{t+1})^2/2\right\} \tag{9}$$

where c is the normalizing constant (we will allways use c as a symbol for this constant regardless of its real value, because this value is of no special interest).

We assume that the parameter transitional density is

$$f(a_{t+1}|a_t) = c\exp\left\{-\frac{1}{2}(a_{t+1}-a_t)^TR(a_{t+1}-a_t)\right\} \tag{10}$$

where R is nonsingular, positive definite and symmetric.

A flow graph for the system is shown in Fig. 1. In the figure n is a sequence of independent Gaussian variables with zero mean and identical precision μ.

It is rather unusual to add the disturbance before the feedback because this will give correlated disturbances in the output, but this is a possible way to construct an exactly observable system because the state allways can be obtained from measured inputs and outputs.

In this special case the solution will be rather trivial and can be obtained by the recursive application of one equation for the mean and one equation for the precision matrix of the density function at every moment t.

The algorithm is given by the rule:

If $f'_t(a_t)$ is given by

$$f'_t(a_t) = c\exp\left\{-\frac{1}{2}(a_t-\alpha_t)^T\Delta_t(a_t-\alpha_t)\right\} \tag{11}$$

then

$$f'_{t+1}(a_{t+1}) = c\exp\left\{-\frac{1}{2}(a_{t+1}-\alpha_{t+1})^T\Delta_{t+1}(a_{t+1}-\alpha_{t+1})\right\} \tag{12}$$

where

$$\begin{aligned}\Delta_{t+1} &= (\Delta_t^{-1}+R^{-1})^{-1} + \mu x_t x_t^T \\ \alpha_{t+1} &= \alpha_t + \mu(y_{t+1} - x_t^T\alpha_t)\Delta_{t+1}^{-1}x_t\end{aligned} \tag{13}$$

The procedure is structuraly equivalent to the Kalman filter[2] and the recursive computation of the mean α_{t+1} of the density function can be done by the system in Fig. 2.

One should observe that one reason why the algorith is so simple is the fact that the input signal value is included in the state vector. The computations are far more complicated if one has to compute the densities of a_{1t} (the "gain" of the system) and $(a_{2t},\ldots,a_{m+n+2,t})$ (the "time constant" of the system) separately. Sometimes this cannot be avoided, for instance if our main interest is concentrated on the gain or the time constant. In this case we denote $(a_{2t},\ldots,a_{m+n+2,t})$ by a_t and compute the density function $f(b_{t+1}|a_{t+1})f(a_{t+1})$ where $b_{t+1} \hat{=} a_{1,t+1}$.
The algorithm is in this case:

Density function at the moment t

$$f'_t(a_t,b_t) = c\exp\left\{-\frac{1}{2}(a_t-\alpha_t)^TG_t(a_t-\alpha_t)-\frac{1}{2}l_t(b_t-\beta_t(a_t))^2\right\} \tag{14}$$

Predicted density function at the moment t

$$\begin{aligned}\hat{f}_t(a_{t+1},b_{t+1}) = c\exp\Big\{&-\frac{1}{2}(a_{t+1}-\alpha'_{t+1})^TG'_{t+1}(a_{t+1}-\alpha'_{t+1}) - \\ &-\frac{1}{2}l'_{t+1}(b_{t+1}-\beta'_{t+1}(a_{t+1}))^2\Big\}\end{aligned} \tag{15}$$

where

$$\begin{cases}G'_{t+1} = (G_t^{-1} + R^{-1})^{-1} \\ \alpha'_{t+1} = \alpha_t\end{cases} \tag{16a}$$

and

$$\left\{\begin{array}{l} k_t \triangleq \mu u_{t-1}/l_t \cdot x_{t-1} \\ l'_{t+1} = (k_t^T(G_t+R)^{-1}k_t + h^{-1} + l_t^{-1})^{-1} \\ \beta'_{t+1}(a_{t+1}) = \beta_t(\alpha_t) - k_t^T(G_t+R)^{-1}R(a_{t+1}-\alpha_t) \end{array}\right. \qquad (16b)$$

where we have assumed that the precision of $(a_{1t},\dots,a_{m+n+2,t})$ is a matrix with structure

$$\left[\begin{array}{c:c} h & 0 \\ \hdashline 0 & R \end{array}\right]$$

which means that we assume that the transitional densities of a and b are independent.

Density function at the moment t+1

$$f'_{t+1}(a_{t+1},b_{t+1}) = c\exp\left\{ -\frac{1}{2}(a_{t+1}-\alpha_{t+1})^T G_{t+1}(a_{t+1}-\alpha_{t+1}) - \right. \\ \left. -\frac{1}{2}l_{t+1}(b_{t+1}-\beta_{t+1}(a_{t+1}))^2 \right\} \qquad (17)$$

where

$$\begin{array}{l} l_{t+1} = l'_{t+1} + \mu u_t^2 \\ \beta_{t+1}(a_{t+1}) = \beta'_{t+1}(a_{t+1}) + \mu u_t/l_{t+1}[y_{t+1}-a_{t+1}^T x_t - u_t\beta'_{t+1}(a_{t+1})] \\ p_{t+1} \triangleq \mu l'_{t+1}/l_{t+1} \\ s_{t+1} \triangleq (G_t+R)^{-1}Rk_t \\ G_{t+1} = G'_{t+1} + p_{t+1}(x_t+u_t s_{t+1})(x_t+u_t s_{t+1})^T. \\ \alpha_{t+1} - \alpha'_{t+1} = \\ = p_{t+1}[y_{t+1} - (\alpha_t^T x_t + \beta_t(\alpha_t)u_t)]G_{t+1}^{-1}(x_t+u_t s_{t+1}) \end{array} \qquad (18)$$

These computations are not in principle more difficult to do than the computations in (13) but they are of course more time consuming.

The marginal densities for a_t and b_t can easily be obtained from the equations above at every moment t.

4. Nonlinear timevarying systems

We start from the assumption that the state transitional density of the nonlinear system is the density function (9) if we assume that there exists an unknown function $d_t: X \rightarrow A_t$ at every moment t such that in (9)

$$a_{t+1} = d_{t+1}(x^{t+1}) , \qquad (19)$$

which means that the value of the parameter is a time varying nonlinear (or linear) function of past and current states.

If this function is completely known and easy to handle we have no problem, but we assume that we have very little information about how the values of the parameter depend on the values of the state and hence we are forced to consider the more general case where we do not make the assumption in equation (19) but have to assume that the parameter

$$(a_o, a_1, \ldots, a_T) = (d_o(x^o), d_1(x^1), \ldots, d_T(x))$$

is a realization of a stochastic variable a.

This means that we have to use our information about the effects of the input and the output on the parameter in some other way.

If we know that the amplitudes of the input signal and output signal vary rather slowly and that because of this the values of the parameter vary slowly, but perhaps rather unpredictably, then we can assume that the parameter is a realization of a Markov parameter with transitional density function (10). The only thing to do after assuming the Markov property is to make a clever guess concerning the optimal value of the precision matrix R in (10). This is of cource one of the main difficulties in applying the tracking algorithm to this type of nonlinear systems and it is a problem which has to be solved (by trial and error) separately in every application.

To test the behaviour of the method simulations were made on two very simple nonlinear systems. The first system was one with exponential nonlinearity and the difference equation for the system can be written in the general form

$$y_t = (a_1+a_2e^{-a_3u_t+a_4})u_t + (b_1+b_2e^{-b_3u_t+b_4})u_{t-1} + n_t \qquad (20)$$

where n_t is a value of a noise signal.

The second system was a saturating system and its difference equation was

$$y_t = \min(a_1, a_2/u_t)u_t + \min(b_1, b_2/u_{t-1})u_{t-1} + n_t . \qquad (21)$$

All simulations were made on the time interval /0,200/ and the input to both systems was

$$u_t = 5(1+\sin(0.015t)) + 0.5\sin(0.5t). \qquad (22)$$

Results of typical simulations are shown in Fig. 3 and Fig. 4.

Fig. 3 shows the result of a simulation of the system with exponential nonlinearity when as parameter values were chosen

$$\begin{array}{ll} a_1 = 0 & b_1 = 0.5 \\ a_2 = 2 & b_2 = 2 \\ a_3 = 0.2 & b_3 = 0.15 \\ a_4 = 0.9 & b_4 = 0.7 . \end{array}$$

The variance of the measurement error n_t was in this simulation 1.5.

Fig. 4 shows the result of a simulation of the saturating system when as parameter values were chosen

$$\begin{array}{ll} a_1 = 2 & b_1 = 1 \\ a_2 = 14 & b_2 = 7.5 . \end{array}$$

The variance of the measurement error n_t was 0.9 in this simulation.

The tracking behaviour is quite acceptable although the constant delay cannot be eliminated with a Markov-1 assumption. In Fig. 4 can be seen the effect of an incorrectly chosen precision matrix. The tracking of the parameter a becomes too slow and the tracking of b too fast which causes overshooting in the tracking of b. The effect of this can be seen in the computed output. In the interval /30,110/ the output is too large and in the interval /140,200/ it is too small.

The correction needed in the value of the precision matrix can easily be done as long as the effect of the error can be seen in the computed output and hence it would be an easy task to repeat the simulation of the saturating system with better result than the one shown in Fig. 4.

5. Conclusions

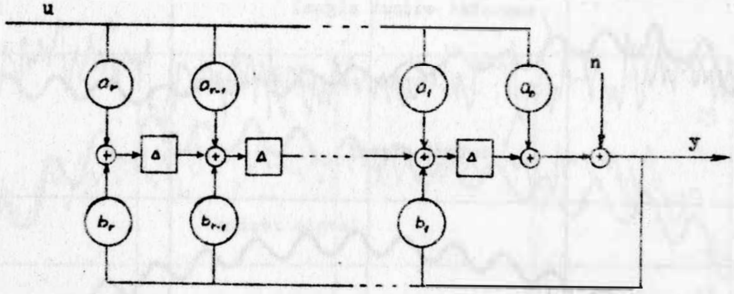
It is obvious that there is no sense in applying the method to systems with known nonlinear characteristics because it is easier and safer to compute all parameter values from input and output signal values if it is possible. The method we have used is simple and efficient in problems where the nonlinear characteristics of the system are unknown or very difficult to use in computations.

Simulations have shown that the method gives satisfactory results even when the precision of the densities is approximated by a constant matrix. If the maximum value of the variance of the solution has been tested by means of the input (for instance using the method outlined in ref. 3) it is quite safe to choose a suitable approximation for the precision.

The algorithm presented in this paper has in a modified form been applied to on-line calibration control of instruments used for measuring process data in digital process control systems, and work on this problem is still being done, but the preliminary results which have been obtained are rather satisfactory and indicate that the method can be used as a means to overcome some of the difficulties caused by slowly varying operating points in nonlinesr systems.

6. References

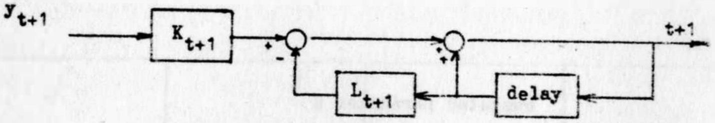
1. Deutsch R.: Estimation Theory, Prentice-Hall 1965.
2. Lee R.C.K.: Optimal Estimation, Identification and Control, M.I.T. Press 1964.
3. Segerståhl B.: Tracking of Parameters in a Timevarying System, Helsinki Technical University, Automatic Control Laboratory. Internal report C 17, 1968.



$r = \max(m,n)$ if $n < m$ $b_{n+1} = \ldots = b_r = 0$

if $m < n$ $a_{m+1} = \ldots = a_r = 0$ Δ = delay

Fig. 1



$$K_{t+1} = \mu \Delta^{-1}_{t+1} x_t$$

$$L_{t+1} = K_{t+1} x^T_t$$

Fig. 2

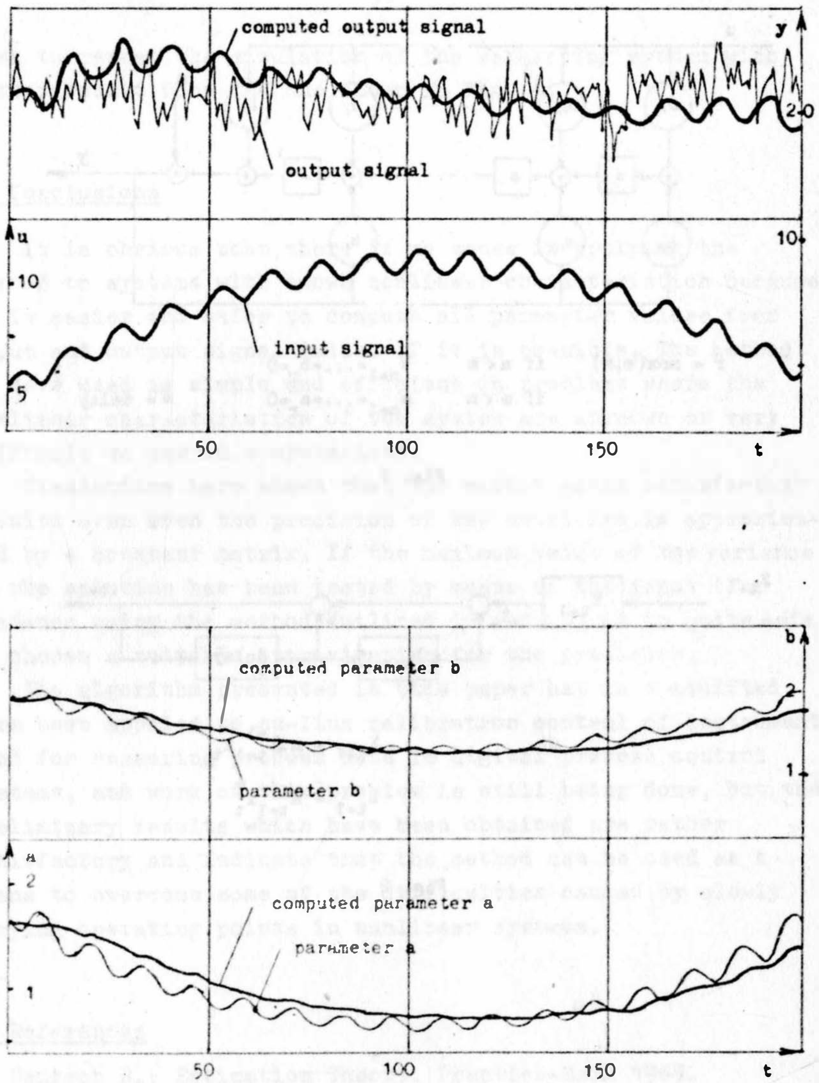

computed output signal
y
2.0
output signal
u
10
10
5
input signal
50
100
150
t
computed parameter b
b
2
1
parameter b
a
2
computed parameter a
parameter a
1
50
100
150
t

Fig. 3

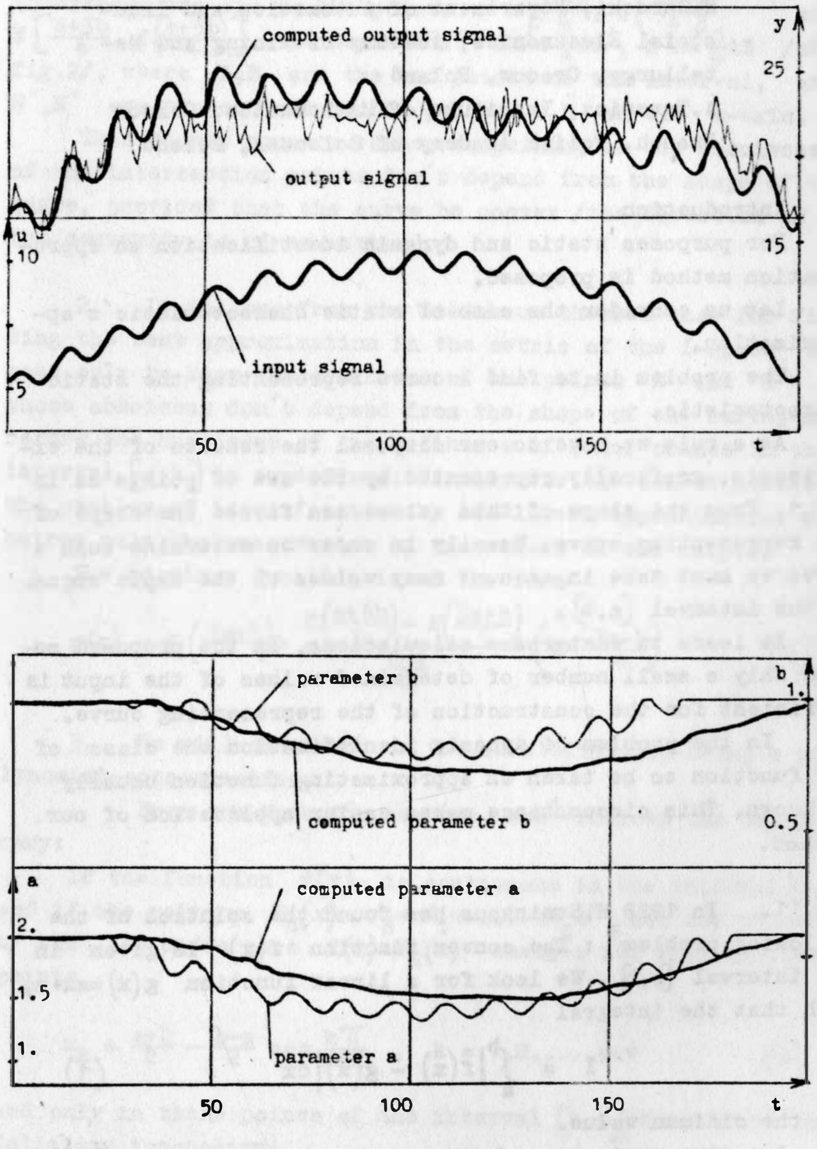

computed output signal
y
25
output signal
u
10
15
input signal
5
50
100
150
t
parameter b
b
1.
computed parameter b
0.5
a
computed parameter a
2.
1.5
1.
parameter a
50
100
150
t

Fig. 4

THE APPROXIMATION METHOD OF IDENTIFICATION

H.Góreoki, Department of Automatios and Industrial Electronios, Academy of Mining and Metallurgy, Craoow, Poland
A.Turowioz, Institute of Mathematios, Craoow Branoh, Polish Academy of Soienoes, Poland

Introduotion

For purposes statio and dynamio identifioation an approximation method is proposed.

Let us consider the oase of statio oharaoteristio's approximation.

The problem is to find a ourve representing the statio charaoteristio.

As a rule we have to our disposal the results of the experiments, grafioally represented by the set of points as in fig.1, From the shape of this set we oan forsee the shape of the representing ourve. Usually in order to determine suoh a ourve we must take in aooount many values of the input signal in the interval $[a,b]$.

It leads to cumbersome oaloulations, In the proposed method only a small number of determined values of the input is suffioient for the construotion of the representing ourve.

In the problem of dynamio identifioation the olass of the funotion to be taken as approximating funotion usually is known. This oiroumstanoe makes easier applioation of our method.

1. In 1928 H.Steinhaus has found the solution of the following problem[3] : The oonvex funotion $f(x)$ is given in the interval $[a,b]$. We look for a linear funotion $g(x)=mx+n$, suoh that the integral

$$I = \int_a^b |f(x) - g(x)| dx \qquad (1)$$

has the minimum value.

In other words we seek the seoant of the oonvex aro suoh that sum of areas lying between the aro and the seoant be minimum. The solution of H.Steinhaus is : The requested secant intersects the arc in the points $M\left[\frac{3a+b}{4}, f\left(\frac{3a+b}{4}\right)\right]$ and $N\left[\frac{a+3b}{4}, f\left(\frac{a+3b}{4}\right)\right]$ it means that $AM' = \frac{1}{4}AB$, $N'B = \frac{1}{4}AB$ /see fig.2/, where A,B are the end points of the interval, and M',N' are projections of the points M,N on the x-axis.

This result appears rather surprising as the abscissas of the intersection points don´t depend from the shape of the curve, provided that the curve be convex /the direction of the convexity is of no matter/.

2. It follows from the Steinhaus theorem that for finding the best approximation in the metric of the L-space we need only to know the ordinates of the points M and N , whose abscissas don´t depend from the shape of the curve, therefore the knowledge of the ordinates of other points in the interval $[a,b]$ is superflous. Finally for the convex curves the problem of identification by the linear approximation requires only the measurement of two values of the output.

The requires function is of the form

$$g(x) = f\left(\frac{3a+b}{4}\right) + \frac{f\left(\frac{a+3b}{4}\right) - f\left(\frac{3a+b}{4}\right)}{\frac{b-a}{2}}\left(x - \frac{3a+b}{4}\right) \qquad (2)$$

3. We will consider now the identification using a polynomial-approximation.

From the Markof´s theorem, see[1] follows the collorary:

If the function $f(x)$ is continuous in the interval $[a,b]$ and if the polynom $P_n(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n$ has the property that the difference $f(x) - P_n(x)$ changes its sign in the points

$$x_k = \frac{a+b}{2} - \frac{b-a}{2}\cos\frac{k\pi}{n+2} \quad , \quad k = 1,2,\ldots,n+1 \qquad (3)$$

and only in these points of the interval $[a,b]$ then holds the following inequality:

$$\int_a^b \left| f(x) - P_n(x) \right| dx \leqslant \int_a^b \left| f(x) - Q_n(x) \right| dx \qquad (4)$$

where $Q_n(x)$ is an arbitraly polynom of the at most n-degree.

The equality in (4) takes place only for $Q_n(x) \equiv P_n(x)$. We remark that

$$P_n(x_k) = f(x_k) \quad , \quad k = 1,\dots,n+1 \tag{5}$$

so the polynom $P_n(x)$ is determined, when we know the values $f(x_1),\dots,f(x_{n+1})$.

In order that the polynom $P_n(x)$ satisfying to (5) changes its sign in the points (3) and only in these points we have a sufficient but not necessary condition: The polynom has the derivative of order $(n+1)$ not vanishing in any point of interval $[a,b]$, see[1]. Therefore if we know the ordinates in points (3) and determine the interpolation-polynom $P_n(x)$ then we obtain the optimal approximation of the function $f(x)$ in the L-space metric, provided that the assumptions of the quoted theorem are fulfilled.

If not, we have an approximation, but we don't know whether it is optimal or not.

Denoting

$$\omega(x) = (x-x_1)(x-x_2)\dots(x-x_{n+1}) \tag{6}$$

we obtain form the Lagrange interpolation formula

$$P_n(x) = \sum_{k=1}^{n+1} f(x_k)\frac{\omega(x)}{(x-x_k)\,\omega'(x_k)} \tag{7}$$

We consider now particular case, see fig.3. Suppose we know that the function $f(x)$ is convex and we want to approximate it with the parabola of the second degree. If we know that the sign of $f^{(3)}(x)$ is constant in the interval $[a,b]$, then the interpolation polynom of the second degree whose interpolation knotes are

$$x_1 = \frac{a+b}{2} - \frac{b-a}{2}\frac{\sqrt{2}}{2}\ , \ x_2 = \frac{a+b}{2}\ , \ x_3 = \frac{a+b}{2} + \frac{b-a}{2}\frac{\sqrt{2}}{2}$$

gives the best approximation.

If the assumptions are not satisfied there is no knowing whether the best approximation is given by a parabola intersecting the arc of the curve in two, three or four points.

Polynom $P_3(x)$ has the form

$$P_3(x)=\frac{1}{(b-a)^2}\Bigg\{\Big[4f(x_1)-8f(x_2)+4f(x_3)\Big]x^2-$$
$$-\Big[f(x_1)\Big(4(a+b)+\sqrt{2}(b-a)\Big)-8f(x_2)(a+b)+$$
$$+f(x_3)\Big(4(a+b)-\sqrt{2}(b-a)\Big)\quad x+\Big[f(x_1)\Big((a+b)^2+\frac{b^2-a^2}{\sqrt{2}}-$$
$$-f(x_2)(a^2+6ab+b^2)+f(x_3)\Big((a+b)^2-\frac{b^2-a^2}{\sqrt{2}}\Big)\Big]\Bigg\} \qquad (8)$$

5. If the function $f(x)$ is known in the whole interval $[a,b]$, then the following formula holds[1]:

$$\int_a^b\left|f(x)-P_n(x)\right|dx=\left|\int_0^{\pi}f\Big(\frac{a+b}{2}+\frac{b-a}{2}\cos\varphi\Big)\operatorname{sgn}\Big[\sin(n+2)\varphi\Big]d\varphi\right| \qquad (9)$$

Due this formula we can calculate the error of the optimal approximation in the L-space metric.

6. We consider now the case of the convex surface

$$z = f(x,y) \qquad (10)$$

where the function $f(x,y)$ is defined in a plane bounded and convex domain G. We may approximate the function (10) by the linear function

$$z = g(x,y) = mx + ny + p \qquad (11)$$

such that the integral

$$I = \iint_G \left| f(x,y)-g(x,y)\right| dx\,dy = F(m,n,p) \qquad (12)$$

attains its minimum value.

We denote by K the intersection curve of the surface (10) with the plane (11), and by L the projection of K on the plane x,y, and finally by H_1 the plane domain bounded by L, see fig.4, and by H_2 the part of domain G exterior to the curve L .

Then we have

$$I = \iint_{H_1} \left[g(x,y) - f(x,y) \right] dx\, dy + \iint_{H_2} \left[f(x,y) - g(x,y) \right] dx\, dy \tag{13}$$

Remark. The equality (13) holds when the convexity of the surface is directed down. In the case of the opposite direction of the convexity we put in (13) (-I) instead of I .

The sign of I is of no matter on the further considerations. Suppose, that $f(x,y)$ has continuous partial derivatives. Then the curve L has in every point a determined normal, and the integral (13) has the partial derivatives with respect to parameters m,n,p . We calculate these derivatives according to Sobolev's formula[2].

In the case of a double integral this formula takes a form : If

$$Q = \iint_{G(t)} \varphi(x,y,t)\, dx\, dy \tag{14}$$

the

$$\frac{dG}{dt} = \iint_{G(t)} \frac{\partial \varphi}{\partial t}\, dx\, dy + \int_{L(t)} \varphi(x,y,t)\, v_n(x,y)\, ds \tag{15}$$

where we denote by $G(t)$ and $L(t)$ the plane domain of integration and its boudary depending on t , and by $v_n(x,y)$ the velocity in the direction of the exterior normal of the curve $L(t)$ in the point (x,y) of this curve, and finally with ds element of the arc of the curve $L(t)$, see fig. 5. If the equation of the curve $L(t)$ is:

$$h(x,y,t) = 0 \tag{16}$$

and the domain $G(t)$ is defined by the inequality

$$h(x,y,t) \geqslant 0 \tag{17}$$

then the gradient is directed into the interior of the domain $G(t)$, and

$$v_n(x,y) = -\frac{\frac{\partial h}{\partial x}\frac{dx}{dt} + \frac{\partial h}{\partial y}\frac{dy}{dt}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}} \qquad (18)$$

Differentiating (16) we obtain:

$$\frac{\partial h}{\partial x}\frac{dx}{dt} + \frac{\partial h}{\partial y}\frac{dy}{dt} + \frac{\partial h}{\partial t} = 0 \qquad (19)$$

hence

$$v_n(x,y) = \frac{\frac{\partial h}{\partial t}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}} \qquad (20)$$

Finally the formula (15) takes a form

$$\frac{\partial Q}{\partial t} = \iint\limits_{h(x,y,t) \geqslant 0} \frac{\partial \varphi}{\partial t}\, dx\, dy + \int\limits_{h(x,y,t) = 0} \varphi(x,y,t) \frac{\frac{\partial h}{\partial t}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}\, ds \qquad (21)$$

Applying Sobolev´s formula to the integral (13) we find

$$\frac{\partial I}{\partial m} = \iint\limits_{H_1} x\, dx\, dy + \int\limits_{L} \left[f(x,y) - g(x,y)\right] \frac{\frac{\partial h}{\partial m}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}\, ds -$$

$$- \iint\limits_{H_2} x\, dx\, dy + \int\limits_{L} \left[g(x,y) - f(x,y)\right] \frac{\frac{\partial h}{\partial m}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}} \qquad (22)$$

$$\frac{\partial I}{\partial n} = \iint\limits_{H_1} y\, dx\, dy + \int\limits_{L} \left[f(x,y) - g(x,y)\right] \frac{\frac{\partial h}{\partial n}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}\, ds -$$

$$- \iint\limits_{H_2} y\, dx\, dy + \int\limits_{L} \left[g(x,y) - f(x,y)\right] \frac{\frac{\partial h}{\partial n}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}\, ds \qquad (23)$$

$$\frac{\partial I}{\partial p} = \iint\limits_{H_1} dx\,dy + \int\limits_L \left[f(x,y) - g(x,y)\right] \frac{\frac{\partial h}{\partial p}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}\, ds -$$

$$- \iint\limits_{H_2} dx\,dy + \int\limits_L \left[g(x,y) - f(x,y)\right] \frac{\frac{\partial h}{\partial p}}{\sqrt{\left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}\, ds \qquad (24)$$

Along the curve L holds the equality

$$f(x,y) = g(x,y) \qquad (25)$$

In consequence the all curvilinear integrals in formulas (22), (23), (24) are equal zero.

Equating to zero the derivatives $\frac{\partial I}{\partial m}$, $\frac{\partial I}{\partial n}$, $\frac{\partial I}{\partial p}$. we obtain

$$\iint\limits_{H_1} x\,dx\,dy = \iint\limits_{H_2} x\,dx\,dy \qquad (26)$$

$$\iint\limits_{H_1} y\,dx\,dy = \iint\limits_{H_2} y\,dx\,dy \qquad (27)$$

$$\iint\limits_{H_1} dx\,dy = \iint\limits_{H_2} dx\,dy \qquad (28)$$

Since

$$\left.\begin{aligned} \iint\limits_{H_1} x\,dx\,dy + \iint\limits_{H_2} x\,dx\,dy &= \iint\limits_{G} x\,dx\,dy \\ \iint\limits_{H_1} y\,dx\,dy + \iint\limits_{H_2} y\,dx\,dy &= \iint\limits_{G} y\,dx\,dy \\ \iint\limits_{H_1} dx\,dy + \iint\limits_{H_2} dx\,dy &= \iint\limits_{G} dx\,dy \end{aligned}\right\} \qquad (29)$$

we have from (26), (27) and (28)

$$\left.\begin{aligned} \iint\limits_{H_1} x\,dx\,dy &= \frac{1}{2}\iint\limits_{G} x\,dx\,dy \\ \iint\limits_{H_1} y\,dx\,dy &= \frac{1}{2}\iint\limits_{G} y\,dx\,dy \\ \iint\limits_{H_1} dx\,dy &= \frac{1}{2}\iint\limits_{G} dx\,dy \end{aligned}\right\} \qquad (30)$$

The formulas (30) express the fact that the area of the domain H_1 is equal to half of the area of the domain G, and the center of gravity of the domain H_1 coincides with the center of gravity of the domain G. As the area of the domain H_1 and the coordinates of its center of gravity depend upon the parameters m,n,p hence the equations (30) constitute the system of three equations with three unknowns m,n,p. From this system we determine the optimal approximating plane.

7. Now we consider the general case of the function of n-variables $f(x_1,\ldots,x_n)$. We want to approximate it with the linear function $g(x_1,\ldots,x_n)=\sum\limits_{k=1}^{n} c_k x_k+d$ in such a way that the integral

$$I = \int\limits_{G}\cdots\int |f(x_1,\ldots,x_n)-g(x_1,\ldots,x_n)|\,dx_1,\ldots,dx_n \qquad (31)$$

yields the minimum.

We assume again that $f(x_1,\ldots,x_n)$ is a convex function in a bounded and convex domain G contained in the n-dimensional space.

Then, if $(a_1,\ldots,a_n)$ and $(b_1,\ldots,b_n)$ are twopoints of the domain G, and λ satisfies to $0 \leqslant \lambda \leqslant 1$, we have

$$f(a_1\lambda+b_1(1-\lambda),\ldots,a_n\lambda+b_n(1-\lambda)) \leqslant \lambda f(a_1,\ldots,a_n)+ \\ +(1-\lambda)f(b_1,\ldots,b_n) \qquad (32)$$

For further consideration it is of no matter if we replace the convex function by a concave one.

Sobolev's formula for n-tuple integral is as follows. Let

$$Q = \int\limits_{G(t)}\cdots\int f(x_1,\dots,x_n,t)\,dx_1,\dots,dx_n \tag{33}$$

Then

$$\frac{dQ}{dt} = \int\limits_{G(t)}\cdots\int \frac{\partial f}{\partial t}\,dx_1,\dots,dx_n + \int\limits_{L(t)}\cdots\int f(x_1,\dots,x_n)\cdot v_n(x_1,\dots,x_n)\,ds \tag{34}$$

where $L(t)$ is the $(n-1)$ dimensional hypersurface bounding the domain $G(t)$, $v_n(x_1,\dots,x_n)$ is the external normal component of the velocity of the displacement of the point $(x_1,\dots,x_n)$ with the change of the parameter t. Finally dS is the element of the surface on $L(t)$.

By the analogous reasoning as in the case of two variables we obtain similar results.

The equation

$$f(x_1,\dots,x_n) = \sum_{k=1}^{n} c_k x_k + d \tag{35}$$

defines the hypersurface L, bounding the domain H_1 contained in the domain G.

The hyperplan is optimal when the volume of domain H_1 is equal to the half of the volume G, and the centers of gravity of these domains coincide.

These requirements give us $(n+1)$ equations for determination of the coefficients $c_1,\dots,c_n,d$.

References :

[1] N.I.Achijezer : Lekcji po tieorii approksimacii. Moskwa 1965, Izd. "Nauka", s. 101-102.

[2] S.L.Sobolew : Urawnienija matiematiczeskoj fiziki. Moskwa 1950, Gosizdat Teor.Lit, s.16-19.

[3] H.Steinhaus : Über die Approximation konvexer mittels linearer Funkctionen, Zeitschrift für Angewandte Mathematik und Mechanik, 8, 1928, s. 414-415.

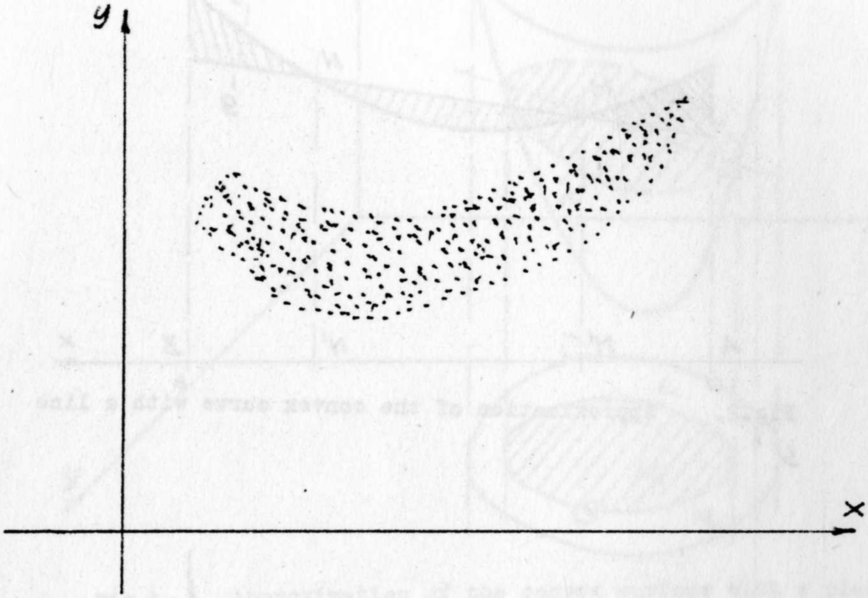


Fig.1. Field of correlation

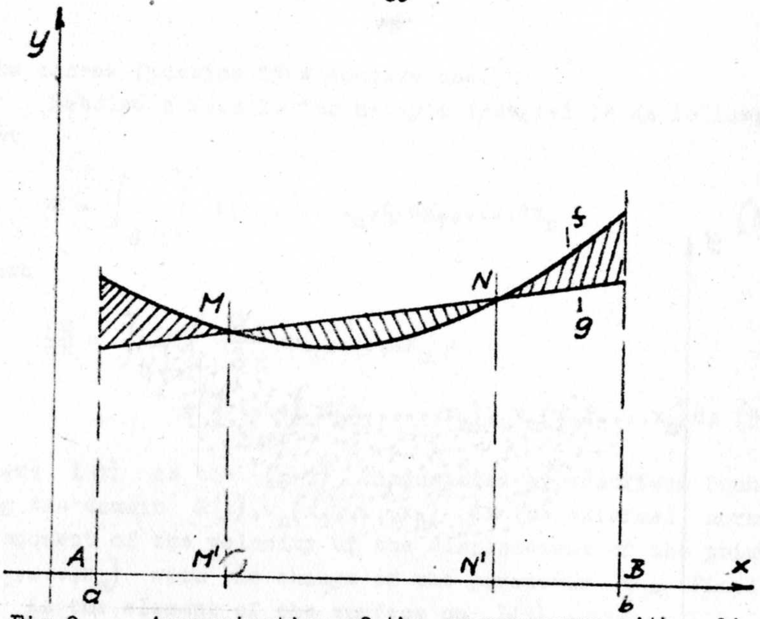


Fig.2. Approximation of the convex curve with a line

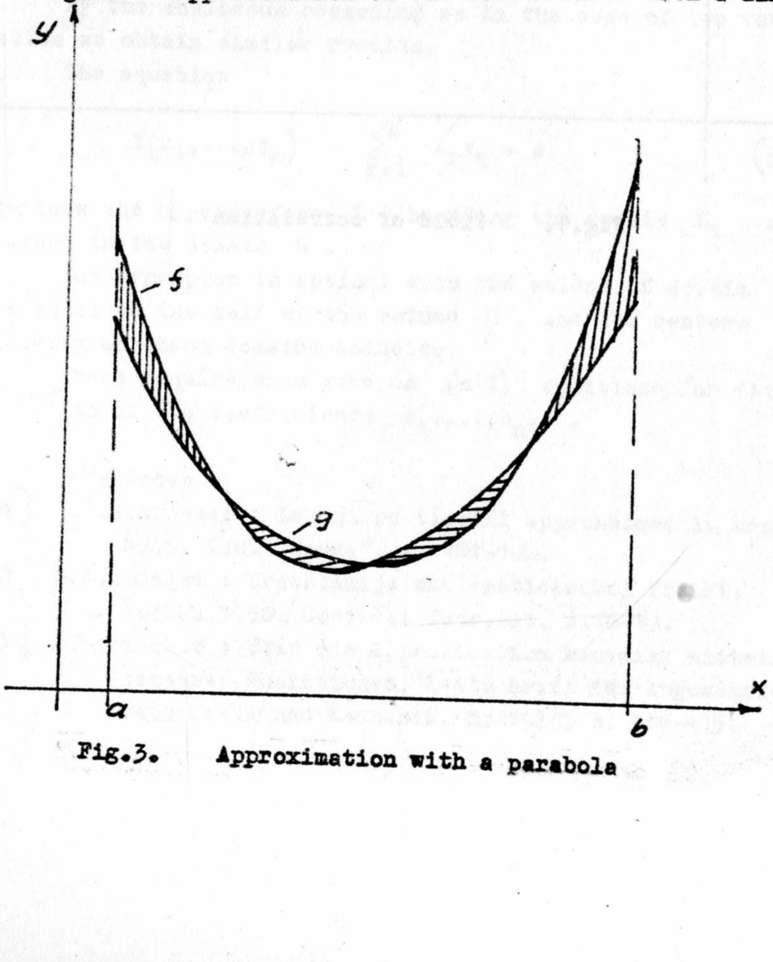


Fig.3. Approximation with a parabola

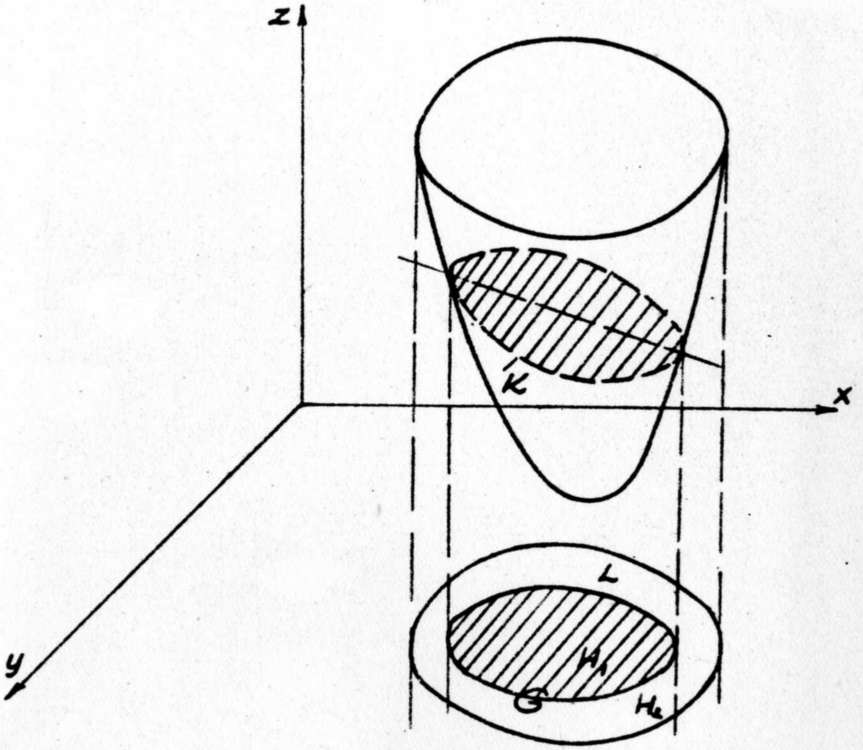


Fig.4. Approximation of the convex surface with a plane

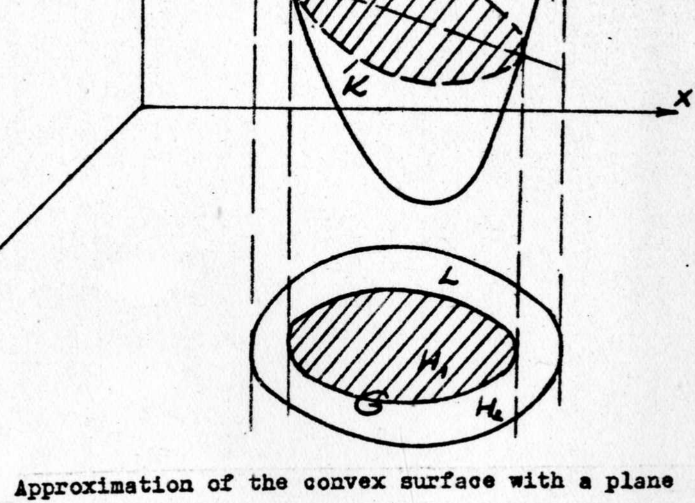


Fig.5. Illustration to Sobolev's formula