

# MODELAÇÃO E DETERMINAÇÃO DE PARÂMETROS CINÉTICOS

FILIPPE GAMA FREIRE

## 1. OBJECTIVO

Durante uma experiência, medem-se certas variáveis, ex.: concentrações, pressões, temperaturas, etc. a que chamaremos  $\vec{y} = [\vec{y}_1, \vec{y}_2, \dots, \vec{y}_m]$ .

Mais ainda, durante cada uma das experiências, há a necessidade de fixar algumas condições iniciais, ex.: a concentração de enzima numa experiência de cinética enzimática, o tempo e a temperatura numa experiência cinética, etc. a que chamaremos  $\vec{x} = [x_1, x_2, \dots, x_n]$ .

Normalmente é necessário escrever um modelo que interprete a experiência. Em geral, o modelo consiste num conjunto de equações baseadas nos princípios da química, física, termodinâmica, cinética e fenómenos de transporte e tenta prever as variáveis  $y$ , que estão a ser medidas. Em geral, as variáveis  $y$  são função de  $x$ . Então, o modelo tem a seguinte forma

$$\vec{y} = \Gamma(\vec{x}, \vec{k}) + \text{erro}$$

O *erro*, normalmente aleatório, provém das medidas de  $y$  cujo valor verdadeiro é desconhecido.

As equações do modelo contêm parâmetros ajustáveis,  $k$ , de modo a levarem em conta o facto de serem fenomenológicos. Por exemplo, as expressões para as velocidades cinéticas contêm constantes de velocidade (parâmetros) cujo valor é desconhecido e não é possível obter por um princípio fundamental.

A estimativa de parâmetros é uma das etapas do processo de formulação e validação de um modelo matemático que descreve um processo em estudo.

- Estimativa de parâmetros clássica - Ou ajuste (linear, não linear)

A estimativa de parâmetros é o processo de obter o valor de parâmetros que ajustam os valores preditos pelo modelo aos valores medidos (dados).

A determinação de parâmetros cinéticos que temos feito até aqui, cai nesta classe.

- Identificação ou Caixa Negra

Neste caso, constrói-se um modelo baseado apenas nos dados de entrada/saída. Este caso aparece quando não se tem nenhum conhecimento a priori do processo em causa.

- Estimativa de parâmetros - Ou modelação (linear, não linear)

Esta classe é muito parecida com a primeira, distingue-se pelo facto do modelo poder não ser uma equação ou conjunto de equações explícitas mas sim poder ser

descrito por um conjunto de equações algébricas e/ou diferenciais não necessariamente solúveis por métodos analíticos ou algébricos, embora também se possa aplicar da mesma forma aos outros casos.

- Ajuste de Curvas

A obtenção de uma função  $f(x)$  que descreva um conjunto de dados  $(x_i, y_i)$ , normalmente com o intuito de interpolar, filtrar ou meramente por questões visuais, é chamado de curve fitting. Os parâmetros que definem a curva são geralmente calculados por minimização de uma medida da qualidade do ajuste, por exemplo os desvios quadrados. Os parâmetros não têm significado físico.

Este método já foi por nós utilizado aquando do método diferencial para se obter os valores da velocidade em função da concentração por diferenciação da curva de ajuste em vez de diferenciação dos resultados.

É objectivo desta pequena introdução focarmo-nos no problema de estimar parâmetros ajustáveis em modelos não lineares descritos por equações algébricas ou equações diferenciais ordinárias.

Os modelos descrevem processos e deste modo explicam o comportamento das variáveis observadas. Assume-se que a estrutura do modelo é conhecida.

São estimados os melhores parâmetros de modo a poderem ser usados na predição para outras condições para as quais é dito que o modelo descreve o comportamento do processo.

Os parâmetros do modelo desconhecidos serão obtidos pela minimização de uma função objectivo adequada. A função objectivo é uma medida da discrepância do modelo aos dados de partida i.e., qualidade do ajuste. Assim, este problema pode também ser visto como um problema de optimização e em princípio podem ser aplicados uma variedade de métodos disponíveis para o resolver.

Enquanto que os engenheiros chamam a este problema "estimativa de parâmetros" os estatístas usam termos como regressão linear ou não linear.

## 2. ESTIMATIVA DE PARÂMETROS PARA CINÉTICA QUÍMICA.

A modelação consiste num conjunto de problemas que se devem resolver mais ou menos sequencialmente, nomeadamente:

- (1) Estrutura do Modelo (Que modelos podem ser usados? Linear, linearizado ou não linear)
- (2) Selecção da função objectivo (O que é que minimizaremos para estimar os parâmetros?)
- (3) Métodos de solução (Como vamos minimizar a função objectivo?)
- (4) Propriedades estatísticas dos parâmetros estimados (Qual a precisão dos parâmetros obtidos?)
- (5) Propriedades estatísticas do modelo-parâmetros (Dado um determinado erro nos parâmetros, qual é a incerteza da previsão do modelo?)
- (6) Testes de validação do modelo (O modelo é suficientemente bom?)
- (7) Teste de discriminação de modelos (Entre vários modelos possíveis, qual é o melhor?)

- (8) Desenho da experiência (Qual o primeiro conjunto de experiências que se deve fazer? Qual será o conjunto de experiências a realizar de modo a obter o máximo de informação possível com o mínimo de esforço de modo a discriminar modelos e obter precisão dos parâmetros)

O ponto 1 foi apresentado nos parágrafos anteriores, onde se aprendeu como expressar a velocidade de reacção aos seus parâmetros e se deu o balanço de massa de um reactor descontínuo.

O ponto 2, nesta introdução é suficiente falar-se de desvios quadrados já descritos para a regressão linear.

Nesta pequena introdução, vamos então apenas falar de um método de resolução muito simples, disponível em todos os computadores pessoais, a folha de cálculo com solver (Microsoft Excel ou Sun OpenOffice), embora o problema possa ser facilmente resolvido noutras plataformas genéricas de matemática como o MatLab ou Mathematica. Resta saber que existem plataformas especializadas na solução deste tipo de problemas em que a mais conhecida é o GAMS.

Os outros pontos, serão de certeza abordados noutras disciplinas do curso. Talvez acerca do ponto 4, 5, 6 e 7 baste para já referir que o somatório dos desvios quadrados obtido após minimização, os gráficos de ajuste e um pouco de bom senso podem ajudar a resolver esses problemas.

Não existindo grandes diferenças na implementação de processos algébricos ou dos diferenciais, vamos falar apenas dos diferenciais devido ao maior interesse para a cinética química.

Vamos pegar num problema e passo a passo vamos levá-lo até ao fim. O exemplo escolhido é a reacção:



que ocorre num reactor descontínuo em fase líquida, caso já abordado anteriormente e resolvido pela integração analítica das equações diferenciais de primeira ordem.

Parte-se de uma solução contendo apenas A e mede-se A e B ao longo do tempo, isto é:  $C_{A_0} \neq 0$ ,  $C_{B_0} = 0$  e  $C_{D_0} = 0$ , não sendo necessário medir a concentração D porque se pode obter por balanço de massa ( $C_A + C_B + C_D = C_{A_0}$ ).

O balanço de massa ao reactor descontínuo a cada uma das espécies é:

$$(2) \quad \begin{aligned} -\frac{dC_A}{dt} &= k_1 C_A \\ \frac{dC_B}{dt} &= k_1 C_A - k_2 C_B \end{aligned}$$

Como se viu anteriormente, estas equações são integráveis e conduzem a:

$$(3) \quad \begin{aligned} C_A &= C_{A_0} e^{-k_1 t} \\ C_B &= C_{A_0} \frac{k_1}{k_2 - k_1} \left( e^{-k_1 t} - e^{-k_2 t} \right) \end{aligned}$$

para  $k_1 \neq k_2$

Tendo-se obtido  $C_D$  pela invariante  $C_D = C_{A_0} - C_A - C_B$ .

$$(4) \quad C_D = C_{A_0} \left( 1 + \frac{k_1 e^{-k_2 t} - k_2 e^{-k_1 t}}{k_2 - k_1} \right)$$

Os valores das constantes de velocidade podem ser obtidos por ajuste das equações 3 aos dados de concentrações versus de tempo.

Vamos agora considerar que, tal como na maioria de processos químicos mais complexos, não era possível integrar as equações 2 e deste modo obter as equações 3 e 4.

Neste caso, temos de integrar aquelas equações por métodos numéricos. Para tal vamos considerar o método mais simples mas também mais robusto, o método de Euler, embora se convide os alunos a rever outros métodos, nomeadamente os de Runge Kutta.

**2.1. Método de Euler.** O método de Euler pode ser derivado de muitas maneiras, mas talvez a mais expedita será recorrendo à definição de derivada, nomeadamente que:

$$\frac{df(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

O método baseia-se em considerar que se escolhe um  $\Delta x$  suficientemente pequeno que garante a igualdade acima, ou que pelo menos, o erro é suficientemente pequeno que não altera os resultados. Este é um problema bastante importante e será resolvido de forma expedita na parte final deste capítulo.

Considerando então que  $\Delta x$  é suficientemente pequeno, podemos re-escrever a equação acima como:

$$\left( \frac{df(x)}{dx} \right)_i = \frac{f_{i+1} - f_i}{\Delta x}$$

ou melhor

$$(5) \quad \boxed{f_{i+1} = f_i + \left( \frac{df(x)}{dx} \right)_i \Delta x}$$

O método de Euler (5) diz-nos então que conhecido um ponto inicial da função a integrar, podemos conhecer todos os pontos de aí para a frente, bastando para isso somar ao ponto anterior a derivada nesse ponto multiplicada pelo incremento.

**2.2. Construção da folha de cálculo para resolver o modelo.** A primeira parte da construção desta folha encontra-se em:

<http://virtualabs.ist.utl.pt/ERI/cinetica/>

e trata-se do exemplo 2 e pode ser referenciado directamente como

<http://virtualabs.ist.utl.pt/ERI/cinetica/consecuti.xls>

Nesta parte, apenas se escreve o modelo e se resolve pelo método de Euler para um conjunto de constantes  $k_1$  e  $k_2$ . Esta folha é interessante também para estudar a influência do passo ( $\Delta t$ ) no resultado obtido. Faz-se notar que um passo adequado para um dado conjunto de constantes pode não o ser para outro conjunto, nomeadamente neste caso, para valores mais elevados destas.

### 2.3. Como verificar que um determinado passo é suficientemente pequeno.

- (1) faça um gráfico das diversas variáveis integradas em função da variável independente ( a variável que foi discretizada, neste caso t).
- (2) fixe a escalas em x e y.
- (3) reduza o passo para metade.
- (4) se verificou alterações no gráfico, para além da possível diminuição de tempo de integração, reduza de novo o passo para metade.
- (5) se não verificou alterações no gráfico, pode considerar adequado o passo para esse conjunto de parâmetros.

Nota: No caso de ir proceder a uma optimização, deve dividir mais uma vez o passo por dois, uma vez que o solver vai variar os parâmetros e deste modo o passo obtido na receita anterior pode não ser suficientemente pequeno podendo resultar num erro do excel ou pior, obter resultados sem significado sem indicação de erro nenhum.

**2.4. Construção da folha de cálculo para obter os parâmetros com base nos resultados de uma só experiência.** Esta folha tem um problema adicional na sua construção. As equações do modelo foram calculadas para um espaçamento de tempo, geralmente, muito menor com que são adquiridos os pontos experimentais. As folhas de cálculo resolvem este problema de uma forma muito expedita.

=LOOKUP(valor,coluna do modelo com esse valor,coluna com o resultado)

deste modo é fácil construir na folha anterior uma nova tabela com as seguintes colunas:

t	$C_A$	$C_B$	$c_a$	$c_b$
t0(re)	re	re	=LOOKUP(t0(re), \$A\$x:\$A\$y, B\$y:B\$y)	=LOOKUP(t0(re), \$A\$y:\$A\$y, C\$y:C\$y)
t1(re)	re	re	=LOOKUP(t1(re), \$A\$y:\$A\$y, B\$y:B\$y)	=LOOKUP(t1(re), \$A\$y:\$A\$y, C\$y:C\$y)
⋮	⋮	⋮	⋮	⋮

em que :

- : re - resultado experimental
- : t0(re), t1(re) - tempos experimentais
- :  $c_a$  - valor do modelo para o tempo t da variável  $C_A$
- :  $c_b$  - valor do modelo para o tempo t da variável  $C_B$
- : \$A\$y : \$A\$y - coluna com a variável t do modelo
- : \$B\$y : \$B\$y - coluna com a variável  $C_A$  do modelo
- : \$C\$y : \$C\$y - coluna com a variável  $C_C$  do modelo

Para além destas colunas, é necessário mais uma em que se calcule o desvio quadrado entre os valores experimentais e o modelo. Esta coluna terá mais uma linha que a anterior onde se colocará o somatório desta coluna e que será a função objectivo a minimizar em que cada linha terá a forma,

$$= (C_A - c_a)^2 + (C_B - c_b)^2$$

**2.5. Minimização de uma função.** Para este problema existe uma grande quantidade de métodos descritos na literatura, mais ou menos complexos de implementar, contudo

com o aparecimento das folhas de cálculo, apareceu uma forma simples e elegante para realizar optimizações sem necessidade de grandes conhecimentos de matemática ou de utilizar software especializado, o "solver".

Após construção de uma folha de cálculo adequada, o problema, se bem enunciado, resolve-se com dois ou três "cliques" de rato.

- (1) Abrir a janela do Solver (depende do sistema operativo, do programa e da versão deste)

Se não encontrar esta ferramenta pode acontecer que não pediu para ser instalada, recorra ao help do programa para obter a informação de como o fazer.

- (2) Seleccionar o problema

**Minimização**, maximização ou obrigar a função objectivo ser 0

- (3) Seleccionar os parâmetros

Seleccionar a coluna de parâmetros que queremos obter

- (4) Seleccionar a função objectivo

Seleccionar a célula onde escrevemos o somatório dos desvios quadrados

- (5) Solve

Clicar em solve!

### 3. CONCLUSÃO

Fica desta forma concluída a construção da tabela e com esta, a folha necessária para realizar a modelação de uma experiência cinética em que se conhece o mecanismo e se obtiveram os resultados experimentais  $t(re)$ ,  $C_A$  e  $C_B$ . Com esta folha é possível obter  $k_1$  e  $k_2$  desse mecanismo.

Em [http://virtualabs.ist.utl.pt/ERI/solver\\_consecutivas.xls](http://virtualabs.ist.utl.pt/ERI/solver_consecutivas.xls)

pode encontrar uma folha para obter essas constantes para um conjunto de dados experimentais.

No caso de não se conhecer o mecanismo por detrás de uma determinada reacção, pode-se estipular diversas hipóteses (mecanismos em série, paralelo, radicalares etc.) que após se minimizar, aquele que apresentar a menor soma dos desvios quadrados será o mecanismo com mais hipóteses de ser o verdadeiro (discriminação de mecanismos).