

**ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC SƯ PHẠM**



BÀI GIẢNG MÔN

QUY HOẠCH THỰC NGHIỆM

(CÁC PHƯƠNG PHÁP THỐNG KÊ XỬ LÝ SỐ LIỆU
THỰC NGHIỆM)

Người soạn: Giang Thị Kim Liên

Đà Nẵng, 2009

Chương 1. CÁC KHÁI NIỆM CHUNG

1.1. Qui hoạch thực nghiệm - bước phát triển của khoa học thực nghiệm

Nhiều công trình nghiên cứu khoa học công nghệ thường đưa đến giải bài toán cực, tìm điều kiện tối ưu để tiến hành các quá trình hoặc lựa chọn thành phần tối ưu để tiến hành các quá trình hoặc lựa chọn thành phần tối ưu của hệ nhiều phần tử. Chẳng hạn, khi xem xét các quá trình CN hóa học mới, nhiệm vụ nghiên cứu thường là thay đổi nhiệt độ, áp suất và tỉ lệ các chất phản ứng để tìm hiệu suất phản ứng cao nhất, tính toán, lựa chọn giá trị thích hợp nhất của các thông số cấu trúc và động học, nhằm đạt đến chất lượng làm việc và hiệu quả kinh tế cao nhất của quá trình. Những bài toán này thường giải quyết ở các mức độ nghiên cứu các yếu tố ảnh hưởng đến hệ, lập mô hình biểu diễn mối phụ thuộc giữa các phần tử của hệ, điều khiển hệ theo mục đích cho trước, hoặc đưa về trạng thái tối ưu theo những chỉ tiêu đánh giá đã chọn. Thông thường các hệ cần điều khiển và tối ưu rất phức tạp, đối tượng nghiên cứu ngày càng đa dạng hơn, trở thành những hệ thống công kênh với tập hợp lớn các yếu tố ảnh hưởng và chỉ tiêu đánh giá. Mối quan hệ giữa các thành phần trong hệ thống càng không thể mô tả bằng các hàm lý thuyết. Vì vậy, đa số các bài toán cực trị được giải quyết bằng thực nghiệm.

Ngày nay người ta thường đề cập tới phương pháp kết hợp giữa lý thuyết và thực nghiệm. Tùy theo mức độ hiểu biết về cơ chế của quá trình, ý nghĩa của nghiên cứu lý thuyết thường được giới hạn ở tác dụng định hướng ban đầu, hỗ trợ giảm bớt khối lượng công việc, rút ngắn thời gian cho nghiên cứu thực nghiệm. Bên cạnh đó, thực nghiệm có tác dụng trở lại, bổ sung cho kết quả nghiên cứu lý thuyết, xác định rõ hơn cơ chế của hiện tượng.

Vai trò của thực nghiệm càng lớn thì mục tiêu đề ra cho chúng càng cao, vì vậy thực nghiệm cũng có nhu cầu phát triển và trở thành đối tượng nghiên cứu, một ngành khoa học.

Có thể nói, lý thuyết qui hoạch thực nghiệm từ khi ra đời đã thu hút sự quan tâm và nhận được nhiều đóng góp hoàn thiện của các nhà khoa học. Những ưu điểm rõ rệt của phương pháp này so với các thực nghiệm cổ điển là:

- Giảm đáng kể số lượng thí nghiệm cần thiết.

- Hàm lượng thông tin nhiều hơn rõ rệt, nhờ đánh giá được vai trò qua lại giữa các yếu tố và ảnh hưởng của chúng đến hàm mục tiêu. Nhận được mô hình toán học thống kê thực nghiệm theo các tiêu chuẩn thống kê, đánh giá được sai số của quá trình thực nghiệm theo các tiêu chuẩn thống kê cho phép xét ảnh hưởng của các yếu tố với mức độ tin cậy cần thiết.

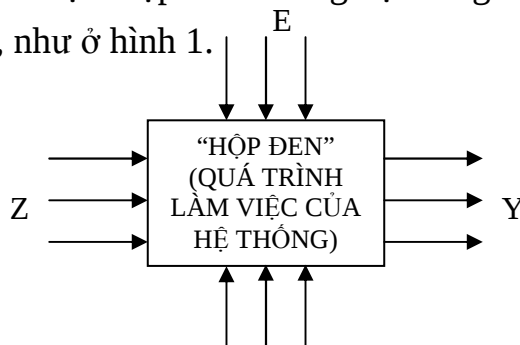
- Cho phép xác định được điều kiện tối ưu đa yếu tố của đối tượng nghiên cứu một cách khá chính xác bằng các công cụ toán học, thay cho cách giải gần đúng, tìm tối ưu cục bộ như các thực nghiệm thụ động.

1.2. Những khái niệm cơ bản của qui hoạch thực nghiệm

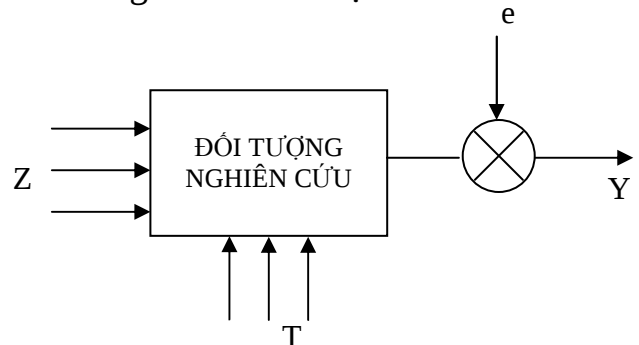
Qui hoạch thực nghiệm là cơ sở phương pháp luận của nghiên cứu thực nghiệm hiện đại. Đó là phương pháp nghiên cứu mới, trong đó công cụ toán học giữa vai trò tích cực. Cơ sở toán học nền tảng của lý thuyết qui hoạch thực nghiệm là toán học xác suất thống kê với hai lĩnh vực quan trọng là phân tích phương sai và phân tích hồi qui.

* **Định nghĩa qui hoạch thực nghiệm:** qui hoạch thực nghiệm là tập hợp các tác động nhằm đưa ra chiến thuật làm thực nghiệm từ giai đoạn đầu đến giai đoạn kết thúc của quá trình nghiên cứu đối tượng (từ nhận thông tin mô phỏng đến việc tạo ra mô hình toán, xác định các điều kiện tối ưu), trong điều kiện đã hoặc chưa hiểu biết đầy đủ về cơ chế của đối tượng.

* **Đối tượng của qui hoạch thực nghiệm trong các ngành công nghệ:** Là một quá trình hoặc hiện tượng nào đó có những tính chất, đặc điểm chưa biết cần nghiên cứu. Người nghiên cứu có thể chưa hiểu biết đầu đủ về đối tượng, nhưng đã có một số thông tin tiên nghiệm dù chỉ là sự liệt kê sơ lược những thông tin biến đổi, ảnh hưởng đến tính chất đối tượng. Có thể hình dung chúng như một “hộp đen” trong hệ thống điều khiển gồm các tín hiệu đầu vào và đầu ra, như ở hình 1.



Hình 1. Sơ đồ đối tượng nghiên cứu



Hình 2. Sơ đồ đối tượng nghiên cứu với nhiễu e có tính cộng

- Các tín hiệu đầu vào được chia thành ba nhóm:

1) Các biến kiểm tra được và điều khiển được, mà người nghiên cứu có thể điều chỉnh theo dự định, biểu diễn bằng vector:

$$Z = [Z_1, Z_2, \dots, Z_k]$$

2) Các biến kiểm tra được nhưng không điều khiển được, biểu diễn bằng vector:

$$T = [T_1, T_2, \dots, T_h]$$

3) Các biến không kiểm tra được và không điều khiển được, biểu diễn bằng vector:

$$E = [E_1, E_2, \dots, E_f]$$

- Các tín hiệu đầu ra dùng để đánh giá đối tượng là vector $Y = (y_1, y_2, \dots, y_q)$. Chúng thường được gọi là các hàm mục tiêu. Biểu diễn hình học của hàm mục tiêu được gọi là mặt đáp ứng (bề mặt biểu diễn).

Phương pháp toán học trong xử lý số liệu từ kế hoạch thực nghiệm là phương pháp thống kê. Vì vậy các mô hình biểu diễn hàm mục tiêu chính là các mô hình thống kê thực nghiệm. Các mô hình này nhận được khi có công tính nhiễu ngẫu nhiên. Cấu trúc mô hình thống kê thực nghiệm có dạng như hình 2.

Trong tập hợp các mô hình thống kê khác nhau, mô hình được quan tâm nhiều nhất trong thực tế là mô hình của phân tích hồi qui. Mô hình hồi qui được biểu diễn bằng quan hệ tổng quát:

$$Y = \varphi(Z_1, Z_2, \dots, Z_k; T_1, T_2, \dots, T_h; \beta_1, \beta_2, \dots, \beta_k) + e = \varphi[(Z, T); \beta] + e$$

Trong đó $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ là vector tham số của mô hình.

Dạng hàm φ được ấn định trước, còn các hệ số β là chưa biết, cần xác định từ thực nghiệm

Để xác định các tham số của mô tả thống kê thực nghiệm ta phải làm các thực nghiệm theo kế hoạch thực nghiệm. Đối tượng nghiên cứu chính của lý thuyết qui hoạch thực nghiệm là các thực nghiệm tích cực. Đó là các thực

thực nghiệm chỉ bao gồm các yếu tố đầu vào thuộc nhóm Z, người thực nghiệm chủ động thay đổi chúng theo kế hoạch thực nghiệm đã vạch sẵn.

*** Các phương pháp qui hoạch thực nghiệm :**

- Thực nghiệm sàng lọc : là thực nghiệm mà nhiệm vụ của nó là tách những yếu tố ảnh hưởng đáng kể ra khỏi những yếu tố đầu vào để tiếp tục nghiên cứu chúng trong các thực nghiệm cần thiết.

- Thực nghiệm mô phỏng : là thực nghiệm liên quan tới việc mô phỏng hiện tượng cần nghiên cứu. Có nhiều dạng mô phỏng, ở đây chỉ quan tâm đến dạng thực nghiệm được hoàn tất bằng mô hình hồi qui đa thức.

- Thực nghiệm cực trị : là thực nghiệm được phát triển từ thực nghiệm mô phỏng. Nhiệm vụ của nó là xây dựng mô hình toán thực nghiệm, theo đó xác định giá trị tối ưu của hàm mục tiêu và các tọa độ tối ưu của hàm. Nói cách khác là xác định bộ kết hợp giá trị các yếu tố mà tại đó hàm mục tiêu đạt cực trị.

*** Kế hoạch thực nghiệm :**

Đối với các thực nghiệm tích cực, miền tác động là miền các giá trị có thể có của các yếu tố Z trong thực nghiệm. Trong miền tác động có miền qui hoạch - miền giá trị của các yếu tố vào Z - trong đó chứa vừa đủ các điểm thí nghiệm của thực nghiệm. Nói cách khác, đó là miền tạo bởi phạm vi thay đổi các yếu tố Z theo kế hoạch thực nghiệm xác định. Kế hoạch thực nghiệm bao gồm các điểm thí nghiệm gọi là điểm của kế hoạch. Đó là một bộ (còn gọi là phương án) kết hợp các giá trị cụ thể của các yếu tố vào Z, ứng với điều kiện tiến hành một thí nghiệm trong tập hợp các thí nghiệm của thực nghiệm. Tại điểm thứ i của kế hoạch, bộ kết hợp các giá trị Z_{ji} bao gồm giá trị cụ thể của k yếu tố đầu vào :

$$Z_{ji} = [Z_{1i}, Z_{2i}, \dots, Z_{ki}]$$

Trong đó: $i = 1, 2, \dots, N$ là điểm thí nghiệm thứ i của kế hoạch thứ N là số điểm thí nghiệm của kế hoạch.

$j = 1, 2, \dots, k$ là yếu tố thứ j ; k là số yếu tố đầu vào.

*** Các mức yếu tố :**

Các giá trị cụ thể của yếu tố vào Z được ấn định tại các điểm kế hoạch gọi là các mức yếu tố. Khái niệm mức yếu tố được sử dụng khi mô tả các điểm đặc trưng trong miền qui hoạch: mức trên, mức dưới, mức cơ sở, mức sao “*”.

Mức cơ sở Z_j^0 của các yếu tố là điều kiện thí nghiệm được quan tâm đặc biệt. Thông thường vector các yếu tố đầu vào tại mức cơ sở $Z^0 = [Z_j^0, Z_j^0, \dots, Z_j^0]$ chỉ ra trong không gian yếu tố một điểm đặc biệt nào đó gọi là tâm kế hoạch, mà trong vùng quanh nó phân bố toàn bộ các điểm kế hoạch. Các tọa độ Z_j^0 của vector Z^0 được chọn theo công thức:

$$X_j = \frac{Z_j - Z_j^0}{\Delta Z_j} ; j = 1, \dots, k$$
$$\Delta Z_j = \frac{Z_j^{\max} - Z_j^{\min}}{2} ; j = 1, \dots, k$$

*** Giá trị mã hóa:** để tiện tính các hệ số thực nghiệm của mô hình hồi qui toán học và tiến hành các bước xử lý số liệu khác, trong kế hoạch thực nghiệm người ta sử dụng các mức yếu tố theo giá trị mã hóa. Giá trị mã hóa của yếu tố là đại lượng không thứ nguyên, qui đổi chuẩn hóa từ các mức giá trị thực của yếu tố nhờ quan hệ :

$$x_j = \frac{Z_j - Z_j^0}{\Delta Z_j} = \frac{2(Z_j - Z_j^0)}{Z_{j\max} - Z_{j\min}}$$

Trong tài liệu này chúng ta giữ nguyên các ký hiệu: Z_j là giá trị thực của yếu tố (gọi là biến thực) ; x_j là giá trị mã hóa của yếu tố (gọi là biến mã).

Như vậy, theo tỉ lệ qui chuẩn, mức cơ sở mã hóa của yếu tố đầu vào là : $x_j^0 = 0$.

Gốc tọa độ của các x_j trùng với tâm thực nghiệm, bước thay đổi của các biến mã x_j ứng với các bước Δx_j chính là 1 đơn vị.

$$\Delta x_j = \frac{Z_{j \max} - Z_{j \min}}{2\Delta Z_j} = 1$$

* ***Ma trận kế hoạch thực nghiệm***: là dạng mô tả chuẩn các điều kiện tiến hành thí nghiệm (các điểm thí nghiệm) theo bảng chữ nhật, mỗi hàng là một thí nghiệm (còn gọi là phương án kết hợp các yếu tố đầu vào), các cột ứng với các yếu tố đầu vào.

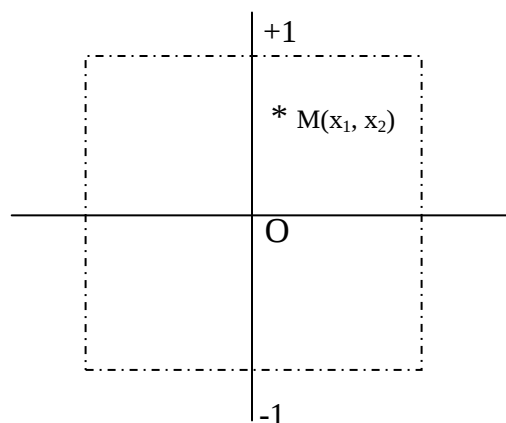
Trong ma trận kế hoạch Z có thể có một số hàng mà mọi thông số vào đều giống nhau, ví dụ, có một số hàng mà mọi thông số vào đều ở mức cơ sở, mọi Z_j^0 .

Ma trận kế hoạch thực nghiệm X là ma trận chỉ gồm toàn các biến mã x_j . Các cột biến mã hoàn toàn khác nhau.

1.3. Các nguyên tắc cơ bản của qui hoạch thực nghiệm

1.3.1. Nguyên tắc không lấy toàn bộ trạng thái đầu vào

Để có thông tin toàn diện về tính chất hàm mục tiêu về nguyên tắc cần tiến hành vô số các thực nghiệm trong miền qui hoạch.



Ví dụ, trong trường hợp có hai yếu tố, nếu cho mỗi yếu tố biến đổi liên tục từ -1 đến +1 thì miền thực nghiệm sẽ là hình vuông chứa vô số điểm $M(x_1, x_2)$ đặc trưng cho trạng thái đầu vào.

Về lý thuyết nếu không tiến hành tất cả các thực nghiệm đó thì có thể bỏ sót đặc điểm nào đó của hàm mục tiêu, tuy nhiên thực tế không thể thực hiện được điều đó. Do vậy người nghiên cứu chỉ có thể lấy những giá trị rời rạc, chọn mức biến đổi nào đó cho các yếu tố. Sự lựa chọn này cần có cơ sở khoa học, nó gắn liền với sự lựa chọn dạng hàm, tức là dạng mô phỏng của bề mặt đáp ứng. Dạng hàm thông thường là bậc một hoặc bậc 2 và số mức biến đổi thường là hai hoặc ba.

1.3.2. Nguyên tắc phức tạp dần mô hình toán học

Khi chưa có thông tin ban đầu về các tính chất của hàm mục tiêu, thì không nên xây dựng mô hình phức tạp của đối tượng để tránh chi phí vô ích về thời gian, phương tiện vật chất nếu không dùng đến mô hình đó. Vì thế lý thuyết qui hoạch thực nghiệm hướng dẫn nên bắt đầu từ những mô hình đơn giản nhất, ứng với những thông tin ban đầu đã có về đối tượng.

Logic tiến hành thực nghiệm là nên làm ít thí nghiệm để có mô hình đơn giản (ví dụ mô hình tuyến tính), kiểm tra tính tương hợp của mô hình :

- Nếu mô hình tương hợp, đạt yêu cầu thì dùng lại, hoặc cải tiến ;
- Nếu mô hình không thì tiến hành giai đoạn tiếp theo của thực nghiệm : làm những thí nghiệm mới, bổ sung để rồi nhận được mô hình phức tạp hơn (ví dụ mô hình phi tuyến), kiểm tra mô hình mới cho đến khi đạt được mô hình hữu dụng.

1.3.3. Nguyên tắc đối chứng với nhiều

Độ chính xác của mô hình phải tương xứng với cường độ nhiễu ngẫu nhiên mà chúng tác động lên kết quả đo hàm mục tiêu. Trong cùng điều kiện như nhau, độ nhiễu càng nhỏ thì mô hình càng phải chính xác, phải phức tạp hơn.

Bằng các công cụ tính toán thống kê, người ta đã xây dựng hoàn chỉnh các qui trình chuẩn theo các tiêu chuẩn thống kê để giải quyết các nhiệm vụ xác

định tính tương hợp của mô hình tìm được, hiệu chỉnh dạng mô hình, kiểm tra tính đúng đắn của các giả thiết, các tiên đề mà dựa vào đó tìm ra các mô hình.

1.4. Các bước qui hoạch thực nghiệm cực trị

1.4.1. Chọn thông số nghiên cứu

Phân loại các yếu tố ảnh hưởng lên đối tượng thành các nhóm Z, T và E. Một mặt đưa ra những biện pháp tích cực để hạn chế tác động của các nhóm yếu tố T và E, mặt khác phải phân tích để chọn từ Z các yếu tố ảnh hưởng chính, loại bớt những yếu tố không cần thiết, nhằm đảm bảo tính khả thi và hiệu quả của thực nghiệm

Lựa chọn chỉ tiêu (mục tiêu) đánh giá đối tượng, sao cho các chỉ tiêu này vừa đáp ứng các yêu cầu của phương pháp qui hoạch thực nghiệm, vừa đại diện nhất cho các điều kiện tối ưu của đối tượng nghiên cứu.

Căn cứ vào số yếu tố ảnh hưởng chính, chỉ tiêu đánh giá, mục đích, nhiệm vụ thực nghiệm, người nghiên cứu phải biết nhóm các yếu tố vào theo kế hoạch thực nghiệm, vì tính hiệu quả và khả năng làm việc của các mô hình hồi qui phụ thuộc nhiều vào kết quả xác định yếu tố vào của chúng.

Trong giai đoạn này, miên qui hoạch và số mức thay đổi của các yếu tố ảnh hưởng phải được xác định sơ bộ.

1.4.2. Lập kế hoạch thực nghiệm

Chọn được dạng kế hoạch thí nghiệm phù hợp với điều kiện tiến hành thí nghiệm và với đặc điểm các yếu tố của đối tượng.

Mỗi dạng kế hoạch đặc trưng bởi các chuẩn tối ưu và tính chất khác nhau. Nên quan tâm nhiều đến điều kiện thí nghiệm và đặc điểm đo đạc, nhận giá trị của mục tiêu.

1.4.3. Tiến hành thí nghiệm nhận thông tin

Sử dụng các phương pháp riêng cho từng đối tượng

Sử dụng một số phương pháp xử lý số liệu, kiểm tra một số giả thiết thống kê. Việc xử lý nhanh các thông tin ngay trong quá trình nhận chúng có tác dụng tích cực, giúp xác minh kịp thời những thí nghiệm cần bổ sung khi điều kiện thí nghiệm còn đang cho phép với các phép kiểm tra đồng nhất phương sai, tính liên thuộc của số liệu bị nghi ngờ, mức độ ảnh hưởng của các yếu tố...

1.4.4. Xây dựng và kiểm tra mô hình thực nghiệm

Sử dụng phương pháp bình phương nhỏ nhất và các nội dung phân tích hồi qui, phân tích phương sai để xác định giá trị của các hệ số trong mô hình hồi qui đa thức, kiểm tra mô hình theo độ tương thích và khả năng làm việc. Tùy theo loại thực nghiệm mà mô hình là tuyến tính hay phi tuyến. Ví dụ các dạng phương trình hồi qui:

- Mô hình bậc hai tuyến tính:

$$y = \varphi(x_1, x_2, \dots, x_k) = b_0 + \sum_{j=1}^k b_j x_j + \sum_{\substack{j,u=1 \\ j \neq u}}^k b_{ju} x_j x_u + \dots$$

- Mô hình bậc hai phi tuyến:

$$y = b_0 + \sum_{j=1}^k b_j x_j + \sum_{\substack{j,u=1 \\ j \neq u}}^k b_{ju} x_j x_u + \dots + \sum_{j=1}^k b_{jj} x_j^2$$

Các hệ số hồi qui $B = [b_0, b_1, b_2, \dots, b_k, b_{11}, b_{12}, \dots, b_{jj}]$ được xác định theo công thức tổng quát dưới dạng ma trận :

$$B = [X^*X]^{-1}X^*Y$$

Trong đó X^* - ma trận chuyển vị của ma trận kế hoạch

Mô hình thống kê thực nghiệm chỉ có thể sử dụng sau khi đã thỏa mãn các tiêu chuẩn thống kê (Student và Fisher).

1.5. Ứng dụng của qui hoạch thực nghiệm trong hóa học, công nghệ hóa học, công nghệ vật liệu và công nghệ môi trường

1.5.1. Thiết lập các mô tả thống kê