

**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.

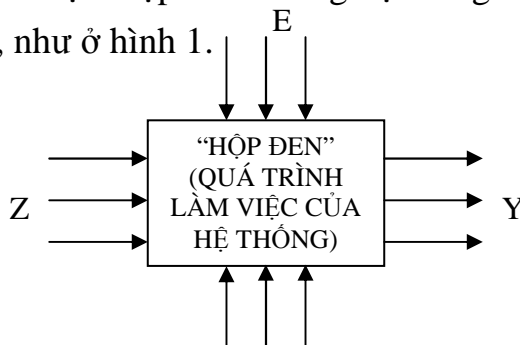
- Cho phép xác định được điều kiện tối ưu đa yếu tố của đối tượng nghiên cứu một cách khá chính xác bằng các công cụ toán học, thay cho cách giải gần đúng, tìm tối ưu cục bộ như các thực nghiệm thụ động.

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

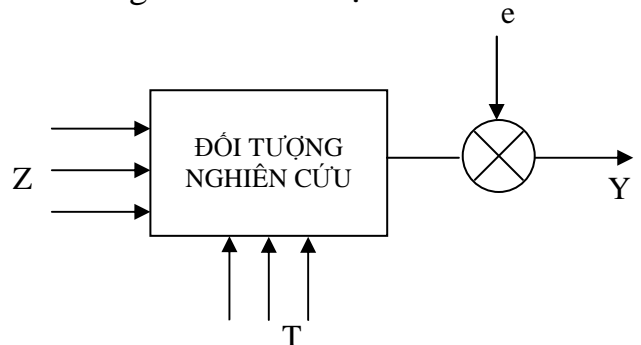
Qui hoạch thực nghiệm là cơ sở phương pháp luận của nghiên cứu thực nghiệm hiện đại. Đó là phương pháp nghiên cứu mới, trong đó công cụ toán học giữa vai trò tích cực. Cơ sở toán học nền tảng của lý thuyết qui hoạch thực nghiệm là toán học xác suất thống kê với hai lĩnh vực quan trọng là phân tích phương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

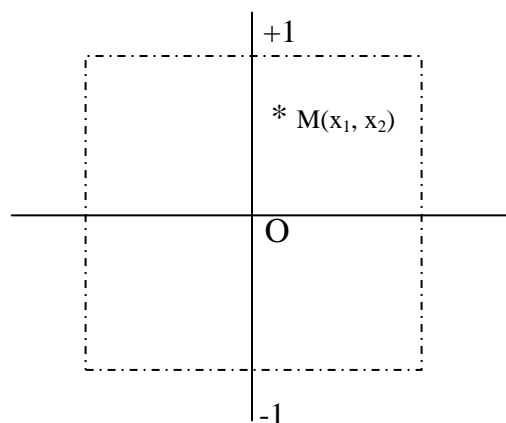
Trong ma trận kế hoạch Z có thể có một số hàng mà mọi thông số vào đều giống nhau, ví dụ, có một số hàng mà mọi thông số vào đều ở mức cơ sở, mọi Z_j^0 .

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 nhiễu

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, miền 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

1) Xác định các yếu tố ảnh hưởng và cấu trúc hệ

Số yếu tố độc lập ảnh hưởng lên quá trình hóa lý bằng số bậc tự do của hệ, được xác định theo công thức :

$$F = F_{dk} + F_h$$

trong đó: F_{dk} là bậc tự do điều khiển

F_h là bậc tự do hình học

Tùy theo yêu cầu của người nghiên cứu mà chỉ cần chọn ra k yếu tố ($k < F$) ảnh hưởng lên một hay nhiều hàm mục tiêu y_q .

Cấu trúc hệ thực hiện quá trình hóa lý : là một hộp đen không biết rõ bản chất bên trong mà chỉ có mối liên hệ bên ngoài giữa hàm mục tiêu và các yếu tố ảnh hưởng.

2) Xác định các hàm toán mô tả hệ

Hàm mô tả hệ là hàm nhiều biến $y = \varphi(x_1, x_2, \dots, x_k)$ được phân tích thành dãy Taylor - hàm hồi qui lý thuyết :

$$y_q = \beta_0 + \sum_{j=1}^k \beta_j x_j + \sum_{j,u=1}^k \beta_{ju} x_j x_u + \dots + \sum_{j=1}^k \beta_{jj} x_j^2$$

Muốn xác định được các hệ số hồi qui lý thuyết β phải cần vô số thí nghiệm. Trong thực tế số thí nghiệm N là hữu hạn, vì vậy mô hình thống kê thực nghiệm có dạng :

$$\hat{y}_q = b_0 + \sum_{j=1}^k b_j x_j + \sum_{\substack{j,u=1 \\ j \neq u}}^k b_{ju} x_j x_u + \dots + \sum_1^k b_{jj} x_j^2$$

Các hệ số b là các tham số của mô tả thống kê.

3) Xác định các tham số mô tả thống kê

Các tham số của mô tả thống kê được xác định từ N thực nghiệm nhờ các kế hoạch thực nghiệm theo phương pháp bình phương cực tiểu. Sau khi tính được các hệ số b phải kiểm tra tính có ý nghĩa của chúng theo tiêu chuẩn Student.

4) Kiểm tra sự tương hợp của mô tả

Sự tương hợp của mô tả thống kê với bức tranh thực nghiệm được kiểm chứng theo tiêu chuẩn Fisher.

1.5.2. Các phương pháp kế hoạch hóa thực nghiệm cực trị chủ yếu

1) Kế hoạch bậc một hai mức tối ưu

Nếu không có thông tin tiên nghiệm cho biết hệ đang ở vùng dừng (vùng phi tuyến, vùng cực trị) thì để mô tả quá trình nên dùng hàm tuyến tính và không có các số hạng bình phương. Để xác định các tham số của nó, nên dùng kế hoạch bậc một hai mức tối ưu của Box-Wilson là kế hoạch toàn phần (2^k) hoặc trong trường hợp cần tiết kiệm thời gian dùng kế hoạch bán phần (2^{k-i}).

2) Kế hoạch bậc hai

Khi mô hình tuyến tính bậc một không tương hợp thì chúng ta là vùng thực nghiệm đã ở vùng phi tuyến, ta phải dùng hàm phi tuyến, có các số hạng bình phương để mô tả.

Có các dạng kế hoạch bậc hai cơ bản :

- Kế hoạch trực giao của Box-Wilson
- Kế hoạch bậc hai tâm xoay của Box - Hunter
- Kế hoạch bậc hai tối ưu của Kiefer

1.6. Khái niệm hệ thống và cách tiếp cận hệ thống công nghệ

Hệ thống: là tập hợp của nhiều phần tử có:

- + Cấu trúc bên trong nhất định.
- + Tương tác với môi trường bên ngoài.

Đề: - Tìm được cấu trúc cần phân tích hệ thành những phần tử
- Nắm được hành vi của hệ phải mô tả tập hợp bản chất của hệ

Vậy nguyên tắc tiếp cận hệ thống: phân tích và tổng hợp → mô tả bản chất của hệ.

Để tìm được bản chất của hệ phải nhờ mô hình hoá và tìm ra được điều kiện công nghệ tối ưu nhờ tối ưu hoá các hàm toán mô tả bản chất của hệ (thường đưa đến giải bài toán cực trị, tức là tìm điều kiện tối ưu để thực hiện một quá trình nhằm đạt đến chất lượng làm việc và hiệu quả kinh tế cao nhất).

1.7 Mô hình hoá

1.7.1. Mô hình

Là một đối tượng được một chủ thể nào đó trên cơ sở của sự đồng dạng về cấu trúc và chức năng dung để thay thế cho một nguyên bản tương ứng để có thể giải quyết một nhiệm vụ nhất định.

Một nguyên bản có thể có nhiều mô hình tùy thuộc vào chủ thể cần giải quyết.

1.7.2. Mô hình toán

Một mô hình toán là biểu diễn toán học những mặt chủ yếu của 1 nguyên bản theo một nhiệm vụ nào đó, trong phạm vi giới hạn với 1 độ chính xác vừa đủ và trong 1 dạng thích hợp cho sự vận dụng.

Một mô hình toán của một nguyên bản phải có 4 điều kiện

- + Chỉ mô tả những mặt chính mà chủ thể quan tâm.
- + Mô tả trong phạm vi giới hạn.
- + Độ chính xác vừa đủ.
- + Khả năng vận dụng mô hình đã được lập trong điều kiện cụ thể.

1.7.3. Các dạng mô hình toán của đối tượng công nghệ hoá học

Xét mô hình thống kê thực nghiệm trong hoá học, CNHH người ta xây dựng quan hệ giữa các đại lượng trên cơ sở thiết lập các quan hệ trên việc xử lý thống kê những giá trị thực nghiệm.

Để xác lập mô tả thống kê của đối tượng CNHH cần thực hiện những bước sau:

- + Xác định số các yếu tố độc lập ảnh hưởng lên hệ, tức là số yếu tố ảnh hưởng (k) lên 1 hay nhiều hàm mục tiêu.
- + Xác định cấu trúc của hệ sẽ được mô hình hoá.

- + Xác định các hàm toán mô tả các quá trình xảy ra trong hệ, và đó thường là hàm nhiều biến và được biểu diễn : $y = f(x_1, x_2, \dots, x_k)$.
- + Xác định các thông số mô hình theo số liệu thực nghiệm.
- + Kiểm tra sự tương thích của mô hình.

1.8. Tối ưu hoá

1.8.1. Khái niệm

Là quá trình tìm kiếm điều kiện tốt nhất (điều kiện tối ưu) của hàm số được nghiên cứu.

Là quá trình xác định cực trị của hàm hay tìm điều kiện tối ưu tương ứng để thực hiện 1 quá trình cho trước.

Để đánh giá điểm tối ưu cần chọn chuẩn tối ưu (là các tiêu chuẩn công nghệ).

1.8.2. Cách biểu diễn bài toán tối ưu

Giả sử một hệ thống công nghệ được biểu diễn dưới dạng sau:

$$Y = F(x_1, x_2, \dots, x_k)$$

x_1, x_2, x_k : k thành phần của vecto thông số đầu vào.

Hàm mục tiêu : $I = I(x_1, x_2, \dots, x_k)$

Bài toán được biểu diễn $I^{opt} = \text{opt } I(x_1, x_2, \dots, x_k) = I(x_1^{opt}, x_2^{opt}, \dots, x_k)$

hoặc $I^{opt} = \max I(x_1, x_2, \dots, x_k)$: đối với bài toán max.

$I^{opt} = \min I(x_1, x_2, \dots, x_k)$: đối với bài toán min.

I^{opt} : hiệu quả tối ưu.

$x_1^{opt}, x_2^{opt}, \dots, x_k$ nghiệm tối ưu hoặc phương án tối ưu.

1.8.3. Thành phần cơ bản của bài toán tối ưu

1.8.3.1. Hàm mục tiêu

- Là hàm phụ thuộc.

- Được lập ra trên cơ sở tiêu chuẩn tối ưu đã được lựa chọn.

→ Hàm mục tiêu là hàm thể hiện kết quả mà người thực hiện phải đạt được

là tiêu chuẩn tối ưu ở dạng hàm, phụ thuộc vào yếu tố đầu vào, giá trị của nó cho phép đánh giá chất lượng của 1 nghiệm cứu.

1.8.3.2. Quan hệ giữa các đại lượng

Các biểu thức toán học mô tả các mối quan hệ giữa tiêu chuẩn tối ưu hoá (hàm mục tiêu) và các thông số ảnh hưởng (thông số cần tối ưu) đến giá trị tiêu chuẩn tối ưu hoá này.

Các quan hệ này thường được biểu diễn bằng phương trình cơ bản hoặc mô hình thống kê thực nghiệm (phương trình hồi qui).

Quan hệ giữa các yếu tố ảnh hưởng với nhau được biểu diễn bằng đẳng thức hoặc bất đẳng thức.

1.8.3.3. Các điều kiện ràng buộc

Để bài toán công nghệ có ý nghĩa thực tế, các biểu thức mô tả điều kiện ràng buộc bao gồm:

- Điều kiện biên.
- Điều kiện ban đầu

Các bước giải bài toán tối ưu:

1. Đặt vấn đề công nghệ : xem xét công nghệ cần được giải quyết là công nghệ gì và chọn ra những yếu tố ảnh hưởng chính

Chỉ ra được hàm mục tiêu $Y : Y \rightarrow \text{MAX}$, hoặc $Y \rightarrow \text{MIN}$

2. Xây dựng mối quan hệ giữa các yếu tố ảnh hưởng và hàm mục tiêu theo qui luật biết trước hoặc mô hình thống kê thực nghiệm

3. Tìm thuật giải: là phương pháp để tìm nghiệm tối ưu của các bài toán công nghệ trên cơ sở các mô tả toán học tương thích đã được thiết lập. Đa số dẫn đến tìm cực trị của các hàm mục tiêu

4. Phân tích và đánh giá kết quả thu được

- Nếu phù hợp \rightarrow kiểm chứng bằng thực nghiệm

- Nếu không phù hợp \rightarrow xem lại từng bước hoặc làm lại từ việc đặt vấn đề

Chương 2. CÁC PHƯƠNG PHÁP PHÂN TÍCH HỒI QUI TƯƠNG QUAN

2.1. Các thông số thực nghiệm

2.1.1. Đại lượng ngẫu nhiên

- Định nghĩa:

Đại lượng ngẫu nhiên (X) là tập hợp tất cả các đại lượng mà giá trị của nó mang lại một cách ngẫu nhiên. Tức là sự xuất hiện là không biết trước.

- Đại lượng ngẫu nhiên X được gọi là rời rạc khi nó nhận hữu hạn hoặc vô hạn các giá trị đếm được khác nhau.

- Đại lượng ngẫu nhiên X được gọi là liên tục nếu nó nhận giá trị bất kì trong một khoảng của trục số.

2.1.2. Sai số đo

Trong thực nghiệm, những giá trị nhận được là giá trị gần đúng của một giá trị thực.

$$\Delta x = x - a \text{ gọi là sai số đo.}$$

Với : a là giá trị thực của một vật.

x là kết quả quan sát được.

Δx là độ lệch giữa a và x.

2.1.2.1. Sai số thô

- Là sai số phạm phải do phá vỡ những điều kiện căn bản của phép đo, dẫn đến các lần đo có kết quả khác nhau nhiều.

- Cách khử sai số thô :

- + Kiểm tra các điều kiện cơ bản có bị vi phạm hay không.

- + Sử dụng một phương pháp đánh giá, để loại bỏ hoặc giữ lại những kết quả không bình thường.

2.1.2.2. Sai số hệ thống

- Là sai số không làm thay đổi trong một loạt phép đo, mà thay đổi theo một quy luật nhất định.

- Nguyên nhân gây sai số: do không điều chỉnh chính xác dụng cụ đo, hoặc một đại lượng luôn thay đổi theo một quy luật nào đó, như nhiệt độ...

- Để khắc phục người ta đặt một hệ số hiệu chỉnh ứng với mỗi nguyên nhân.

2.1.2.3. Sai số ngẫu nhiên

- Sai số ngẫu nhiên của phép đo là đại lượng ngẫu nhiên đặc trưng bằng luật phân phối thể hiện mối quan hệ giữa các giá trị có thể có của sai số và xác suất để sai số ngẫu nhiên nhận các giá trị ấy.

- Là sai số còn lại sau khi đã khử sai số thô và sai số hệ thống.

- Sai số ngẫu nhiên do nhiều yếu tố gây ra, tác dụng rất nhỏ, không thể tách riêng ra, vì thế không loại trừ được.

2.1.3. Các đặc trưng số của đại lượng ngẫu nhiên

2.1.3.1. Kỳ vọng

1. Kỳ vọng toán của biến ngẫu nhiên

- Định nghĩa:

Kỳ vọng toán của biến ngẫu nhiên X là số đặc trưng cho giá trị trung bình tính theo xác suất của tất cả giá trị của X.

Cho X là biến ngẫu nhiên, kỳ vọng toán của biến ngẫu nhiên X được kí hiệu là E(X) và xác định như sau:

- Nếu X là biến ngẫu nhiên rời rạc và giá trị x_i có thể nhận các xác suất p_i

$$(i = 1, 2, \dots) \text{ thì: } E(X) = \sum_{i=1}^n p_i x_i \quad (2.1)$$

- Nếu X là biến ngẫu nhiên liên tục có hàm mật độ xác suất là $f(x)$ thì:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx \quad (2.2)$$

2. Kỳ vọng mẫu thực nghiệm

Kỳ vọng mẫu thực nghiệm được xác định bằng giá trị trung bình của các số liệu quan sát của mỗi phép đo.

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m x_i \quad (2.3)$$

Trong đó: x_i là số đo của đại lượng x ở lần đo thứ i.

m là số lần đo.

3. Mod của biến ngẫu nhiên

Mod của biến ngẫu nhiên rời rạc X là điểm x_0 sao cho:

$$P(X = x_0) = \max P(X = x_i)$$

$i = 1, 2, \dots$, tức là tại đó xác suất x_i là lớn nhất.

2.1.3.2. Phương sai điều chỉnh mẫu thực nghiệm

Phương sai là đặc trưng quan trọng để phản ánh độ phân tán giá trị biến ngẫu nhiên xung quanh kỳ vọng và được kí hiệu là S^2 .

1. Phương sai mẫu thực nghiệm

Giả sử x_1, x_2, \dots, x_m là mẫu thực nghiệm của X , khi đó S^2 gọi là phương sai mẫu thực nghiệm của X , và được xác định như sau:

$$S^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})^2 \quad (2.4)$$

Trong đó: S^2 là phương sai mẫu thực nghiệm.
 m là số lần đo hay số lần quan sát được.
 x_i là số đo của đại lượng x ở lần đo thứ i .
 \bar{x} là trung bình mẫu thực nghiệm.

2. Phương sai điều chỉnh mẫu thực nghiệm

Giả sử S^2 là phương sai mẫu thực nghiệm, khi đó số thực S_1^2 được gọi là phương sai mẫu hiệu chỉnh của X và được xác định như sau:

$$S_1^2 = \frac{1}{f} \sum_{i=1}^m (x_i - \bar{x})^2 \quad (2.5)$$

$f = m - 1$ là bậc tự do đặc trưng cho mẫu thực nghiệm.

2.1.3.3. Độ lệch chuẩn (SD)

- Là tham số dùng để xác định độ phân tán của biến ngẫu nhiên có cùng đơn vị với nó.

- Giả sử S^2 và S_1^2 là phương sai và phương sai điều chỉnh mẫu ngẫu nhiên của X , khi đó S và S_1 được gọi là độ lệch tiêu chuẩn điều chỉnh mẫu thực nghiệm của X và xác định như sau:

$$S = \sqrt{S^2} \quad (2.6)$$

$$S_1 = \sqrt{S_1^2} \quad (2.7)$$

2.1.3.4. Sai số chuẩn (SE)

- Là tỷ lệ giữa độ lệch chuẩn trung bình mẫu với căn bậc hai của dung lượng mẫu:

$$SE = \sigma = \frac{S_1}{\sqrt{N}} \quad (2.8)$$

- Là thông số thống kê quan trọng để đánh giá mức độ phân tán của mẫu chính nó biểu thị sai số của số trung bình. Sai số ở đây do sự chênh lệch cơ

học có hệ thống của số liệu mà phương thức chọn mẫu là một trong những nguyên nhân chính gây nên.

- Mục đích chính SE là xác định mức độ phân tán của giá trị trung bình mẫu và giới hạn tin cậy của mẫu thực nghiệm.

2.1.3.5. Ý nghĩa của phương sai, độ lệch chuẩn, sai số chuẩn

Phương sai, độ lệch chuẩn, sai số chuẩn giúp cho ta nhận biết được mức độ đồng đều của giá trị thực nghiệm.

Nếu phương sai, độ lệch chuẩn, sai số chuẩn nhỏ thì các giá trị thực nghiệm tương đối đồng đều và tập trung xung quanh giá trị trung bình.

2.1.4. Độ chính xác và độ tin cậy của phép đo

- Giả sử một phép đo với sai số tin cậy như sau:

$$|X - \bar{X}| = \Delta X = \mathcal{E}$$

Độ tin cậy γ là xác suất để kết quả các lần đo rơi vào khoảng tin cậy $(\bar{X} - \mathcal{E} < X < \bar{X} + \mathcal{E})$, tức là $P(\bar{X} - \mathcal{E} < X < \bar{X} + \mathcal{E}) = \gamma$ và độ tin cậy thường cho trước 0,95; 0,99; 0,999;...

2.2. Phân tích thống kê các kết quả thực nghiệm (phân tích quy hồi)

Gồm các bước sau:

- Kiểm tra giá trị của tất cả các hệ số hồi qui bằng cách so sánh với sai số lặp lại (S_{bj}) hay còn gọi là sai số chuẩn.

- Sự phù hợp giữa mô tả toán học với kết quả thực nghiệm.

2.2.1. Phương sai tái hiện

Xác định phương sai tái hiện để xác định sai số tái hiện.

2.2.1.1. Phương sai tái hiện của một thí nghiệm

Giả sử một thí nghiệm được lặp đi lặp lại m lần với giá trị tương ứng thu được là y_1, y_2, \dots, y_m . Phương sai tái hiện của một mẫu thực nghiệm được

xác định như sau:

$$S_{th}^2 = \frac{1}{f} \sum_{i=1}^m (y_i - \bar{y})^2 \quad (2.9)$$

$$\text{hay} \quad S_{th}^2 = \frac{1}{m-1} \sum_{i=1}^m (y_i - \bar{y})^2 \quad (2.10)$$

Trong đó: $f = m - 1$ là độ tự do đặc trưng cho khả năng biến đổi mà không làm thay đổi hệ.

m là số lần lặp.

2.2.1.2. Phương sai tái hiện của một cuộc thí nghiệm

$$S_{th}^2 = \frac{1}{N} \sum_{u=1}^N S_u^2 \quad (2.11)$$

$$\text{Mà } S_u^2 = \frac{1}{m-1} \sum_{i=1}^m (y_{ui} - \bar{y}_u)^2 \quad (2.12)$$

$$\Rightarrow S_{th}^2 = \frac{1}{N(m-1)} \sum_{u=1}^N \sum_{i=1}^m (y_{ui} - \bar{y}_u)^2 \quad (2.13)$$

Trong đó: $u = 1, 2, 3, \dots$

$i = 1, 2, 3, \dots$

N là số thí nghiệm khác nhau.

m là số lần lặp lại.

Công thức dùng để tính phương sai tái hiện của một cuộc thí nghiệm, thường sử dụng cho phương án thí nghiệm song song (phương án mà mỗi một điểm thí nghiệm phải tiến hành lặp lại).

Phương sai phân phối trung bình cho từng thí nghiệm được xác định như sau :

$$S_{th}^2(\bar{y}) = \frac{1}{m} S_{th}^2 \quad (2.14)$$

• **Ví dụ 1:** Tính phương sai tái hiện của một cuộc thí nghiệm tương ứng với những số liệu thực nghiệm thu được ở bảng sau :

Bảng 1: Kết quả thí nghiệm

S.T.N (u)	Số lần lặp (m)	Kết quả			
		y_{u1}	y_{u2}	y_{u3}	y_{u4}
1	3	73	69	68	70
2	3	58	58	64	60
3	3	54	59	52	55
4	3	84	94	92	90
5	3	100	106	109	105
6	3	98	90	97	95
7	3	77	85	78	80
8	3	105	95	100	100

Từ bảng số liệu ta thấy $i = 1,2,3$; $u = 1,2,3\dots,8$; $m = 3$; $N = 8$.

Để tính phương sai tái hiện của một cuộc thí nghiệm ta lập bảng sau:

Bảng 2: Phương sai tái hiện của từng thí nghiệm

u	$(y_{u1} - \bar{y}_u)^2$ (1)	$(y_{u2} - \bar{y}_u)^2$ (2)	$(y_{u3} - \bar{y}_u)^2$ (3)	(1+2+3)	S_u^2
1	9	1	4	14	7
2	4	4	16	24	12
3	1	16	9	26	13
4	36	16	4	56	28
5	25	1	16	42	21
6	9	25	4	38	19
7	9	25	4	38	19
8	25	25	0	50	25
					$\sum_{u=1}^8 = 144$

Từ kết quả ở bảng 2, ta tính phương sai tái hiện của cuộc thí nghiệm:

$$S_{th}^2 = \frac{1}{N} \sum_{u=1}^N S_u^2 = \frac{144}{8} = 18$$

Phương sai phân phối trung bình cho một thí nghiệm:

$$S_{th}^2(\bar{y}) = \frac{1}{m} S_{th}^2 = \frac{18}{3} = 6$$

2.2.2. Phương sai dư

- Độ dư là hiệu giữa giá trị thực nghiệm thu được với giá trị tính được theo phương trình hồi qui của các thông số tối ưu.

- Phương sai tìm được trên cơ sở tổng bình phương các độ dư gọi là phương sai dư, được kí hiệu và xác định như sau:

$$S_{dur}^2 = \frac{1}{f_{dur}} m \sum_{u=1}^N (\tilde{y} - \bar{y}_u)^2 \quad (2.15)$$

$$S_{dur}^2 = \frac{1}{f_{dur}} \sum_{u=1}^N (\tilde{y} - \bar{y}_u)^2 \quad (2.16)$$

Trong đó: $f_{dur} = N - L$ độ tự do dư.

N là số thí nghiệm trong một cuộc thí nghiệm.

L số hệ số có nghĩa trong phương trình hồi qui.

\tilde{y}_u giá trị được tính theo phương trình hồi qui ứng với điều kiện thí nghiệm thứ u .

\bar{y}_u là giá trị trung bình thực nghiệm tại thí nghiệm thứ u (trong điều kiện mỗi điểm thực nghiệm được tiến hành lặp lại).

y_u là giá trị thực nghiệm trong điều kiện không làm thí nghiệm lặp.

2.2.3. Kiểm định thống kê

2.2.3.1. Kiểm tra sự đồng nhất của các phương sai

- Kiểm tra sự đồng nhất của các phương sai là kiểm tra độ hội tụ của các giá trị thực nghiệm. Phương pháp kiểm tra này chỉ được áp dụng trong phương án thí nghiệm song song.

- Để kiểm tra người ta chỉ sử dụng chuẩn Cochoren.

Các điểm phân vị của phân phối chuẩn Cochoren với $P = 0,05$

Số T.N	Độ tự do ($f = m - 1$)		
	1	2	3
2	0,9985	0,9750	0,9392
3	0,9669	0,8709	0,7977
“	“	“	“
“	“	“	“
∞	0,000	0,000	0,000

Trong đó: N là số thí nghiệm trong một cuộc thí nghiệm.

f là độ tự do ứng với thí nghiệm có phương sai tái hiện lớn nhất.

m là số lần lặp của thí nghiệm có phương sai tái hiện lớn nhất.

G_b được tìm thấy ở bảng với mức ý nghĩa đã chọn, là điểm gặp nhau giữa hàng biểu thị số thí nghiệm N và cột biểu thị bậc tự do f .

* Các bước tiến hành kiểm tra

- Xác định đại lượng trung bình từ các kết quả của các thí nghiệm song song.

- Xác định các phương sai thực nghiệm (S_u^2) tại mỗi điểm thí nghiệm theo công thức (2.9).

- Tính tổng các phương sai $\sum_{u=1}^N S_u^2$

- Tính G_{tn} theo công thức sau:
$$G_{tn} = \frac{\max S_u^2}{\sum_{u=1}^N S_u^2} ; u = 1,2,3,\dots,N \quad (2.17)$$

$\max S_u^2$ là giá trị cực đại của phương sai thực nghiệm thứ u .

N là số thí nghiệm trong một cuộc thí nghiệm.

- Tra bảng G_b với mức ý nghĩa P đã chọn, số thí nghiệm N và độ tự do f của điểm thực nghiệm có phương sai tái hiện lớn nhất.

- So sánh G_{tn} và G_b .

+ Nếu $G_{tn} < G_b$: giả thiết được chấp nhận.

+ Nếu $G_{tn} > G_b$: giả thiết không được chấp nhận.

2.2.3.2. Kiểm tra ý nghĩa của các hệ số trong phương trình hồi qui

- Mục đích của kiểm tra này là xem các hệ số b_j trong phương trình hồi qui có khác 0 với một độ tin cậy nào đó hay không.

- Để kiểm tra ý nghĩa của các hệ số trong phương trình hồi qui ta phải sử dụng chuẩn Student (t).

* Các bước tiến hành kiểm tra:

- Tính chuẩn t_{tn} theo công thức:
$$t_{tn} = t_j = \frac{|b_j|}{S_{b_j}} \quad (2.18)$$

b_j là hệ số ứng với yếu tố thứ j trong PTHQ; $j = 0,1,2,\dots$

S_{b_j} độ lệch quân phương của hệ số b_j

- Tra bảng $t_b (P,f)$ ứng với mức ý nghĩa P chọn trước và f ; f là bậc tự do ứng với phương sai tái hiện của từng phương án mà người nghiên cứu đã chọn.

- So sánh t_j và t_b

+ Nếu $t_j > t_b$ hệ số b_j có ý nghĩa và được giữ lại trong PTHQ.

+ Nếu $t_j < t_b$ hệ số b_j không có ý nghĩa và loại khỏi PTHQ. Các hệ số còn lại được tính lại theo phương pháp bình phương tối thiểu cho tới khi tất cả chúng đều có nghĩa.

2.2.3.3. Kiểm tra sự tương thích của PTHQ với thực nghiệm

- Dạng PTHQ là do người nghiên cứu tự chọn và các hệ số trong PTHQ được xác định dựa trên các số liệu thực nghiệm. Vì vậy cần phải xem xét mô tả toán học đó có phù hợp với thực nghiệm hay không, và người ta dùng phân phối Fisher (F) với một mức ý nghĩa nào đó.

* Các bước tiến hành kiểm tra:

- Viết PTHQ với các hệ số có nghĩa.

- Tính F_{tn} theo công thức:
$$F_{tn} = \frac{S_u^2}{S_{th}^2} \quad (2.19)$$

Trong đó: S_u^2 là phương sai tương thích và được tính theo công thức (2.15), (2.16).

S_{th}^2 là phương sai tái hiện được tính theo công thức (2.8) với phương án thí nghiệm tại tâm hoặc tính theo công thức (2.14) ứng với phương án thí nghiệm song song.

- F_b tra bảng $f_b(p, f_1, f_2)$ tức là ứng với mức ý nghĩa P đã chọn và bậc tự do f_1, f_2

- Tiêu chuẩn kiểm định (so sánh F_{tn} và F_b)

+ Nếu $F_{tn} < F_b$ thì PTHQ vừa lập phù hợp với thực nghiệm.

+ Nếu $F_{tn} > F_b$ thì PTHQ vừa lập không phù hợp với thực nghiệm và làm tiếp các công việc sau:

* Kiểm tra lại công việc tính toán.

* Xem lại mô hình nghiên cứu đã đúng chưa.

* Chọn mô tả toán học (PTHQ) ở mức cao hơn.

2.3. Các phương pháp phân tích hồi quy

2.3.1. Phương pháp bình phương nhỏ nhất (BPNN)

Là phương án cơ bản có hiệu lực khi xử lý các số liệu thực nghiệm và xây dựng mô hình thống kê cho nhiều đối tượng nghiên cứu thuộc các lĩnh vực khác nhau.

Phương pháp này cho phép xác định các hệ số của phương trình hồi qui đã chọn sao cho độ lệch của sự phụ thuộc đã cho so với số liệu thực nghiệm là nhỏ nhất.

$$\Phi = \sum_{u=1}^N (\tilde{Y}_u - Y_u)^2 \rightarrow \min \quad (2.20)$$

Trong đó: Y_u là giá trị thực nghiệm ứng với k thông số tối ưu ở thí nghiệm thứ u .

\tilde{Y}_u là giá trị theo phương trình hồi qui số tối ưu ở thí nghiệm thứ u .

2.3.2. Hồi quy tuyến tính một biến

Phương trình hồi qui tuyến tính một biến số có dạng:

$$\hat{y} = b_0 + b_1x \quad (2.21)$$

Các hệ số của phương trình hồi quy được xác định bằng phương pháp bình phương nhỏ nhất (BPNN), với số thí nghiệm là N.

Hệ phương trình chuẩn có dạng :

$$\left. \begin{aligned} \sum y_i - \sum (b_0 + b_1x_i) &= 0 \\ \sum y_i x_i - \sum (b_0 + b_1x_i)x_i &= 0 \end{aligned} \right\}$$

2.3.3. Hồi quy parabol

Phương trình hồi quy parabol - bậc hai một biến có dạng:

$$\hat{y} = b_0 + b_1x + b_{11}x^2$$

Các hệ số của phương trình hồi quy cũng được xác định bằng phương pháp bình phương nhỏ nhất (BPNN), với số thí nghiệm là N.

Trong trường hợp này :

$$\frac{\partial f(x)}{\partial b_0} = 1 ; \quad \frac{\partial f(x)}{\partial b_1} = x ; \quad \frac{\partial f(x)}{\partial b_{11}} = x^2$$

Hệ phương trình chuẩn có dạng :

$$\left. \begin{aligned} b_0N + b_1 \sum x_i + b_{11} \sum x_i^2 &= \sum y_i \\ b_0 \sum x_i + b_1 \sum x_i^2 + b_{11} \sum x_i^3 &= \sum x_i y_i \\ b_0 \sum x_i^2 + b_1 \sum x_i^3 + b_{11} \sum x_i^4 &= \sum x_i^2 y_i \end{aligned} \right\}$$

2.3.4. Hồi quy hàm số mũ

Khi số thực nghiệm N bé, nếu tăng bậc của đa thức có thể dẫn đến việc tăng phương sai dư. Lúc này để giảm số các hệ số không xác định, ta dùng hồi quy hàm số mũ. Việc xác định các hệ số của phương trình hồi quy có thể rất khó khăn do phải giải hệ phương trình phi tuyến. Việc tính toán sẽ trở nên đơn giản hơn nếu tiến hành thay thế các biến số và hạ bậc đa thức.

Ví dụ các quan hệ kiểu hàm số mũ như sau :

$$\hat{y} = b_0 b_1^x \qquad \hat{y} = b_0 x^{b_1}$$

Được logarit hóa :

$$\lg \hat{y} = \lg b_0 + x \lg b_1 \qquad \lg \hat{y} = \lg b_0 + b_1 \lg x$$

Sau khi đặt :

$$\lg \hat{y} = \hat{z}; \lg b_1 = a_1; \lg b_0 = a_0; \lg x = t$$

Ta sẽ nhận được phương trình tuyến tính với các biến số :

$$\hat{z} = a_0 + a_1 x \qquad \hat{z} = a_0 + b_1 t$$

Các hệ số a_0, a_1, b_1 được xác định theo phương pháp BPNN. Từ a_0 và a_1 có thể tính được b_0 và b_1 .

2.3.5. Hồi quy nhiều biến

Nếu cần nghiên cứu liên kết tương quan giữa nhiều đại lượng người ta dùng phương trình hồi quy nhiều biến :

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

Ở đây, chúng ta gặp không phải đường hồi quy, mà là mặt phẳng hồi quy khi $k=2$ và mặt hyper khi $k>2$. Trong trường hợp chung, bề mặt này gọi là bề mặt mức hoặc bề mặt đáp trị. Khi xây dựng bề mặt mức trên trục tọa độ của không gian yếu tố cần phải đặt các giá trị bằng số của các yếu tố lên hệ tọa độ. Phải chuyển từ quy mô tự nhiên sang quy mô chuẩn. Nghĩa là phải tiến hành chuẩn hóa tất cả các giá trị của các đại lượng ngẫu nhiên theo các công thức thống kê và chuyển từ biến thực sang biến được mã hóa không có thứ nguyên.

Chương 3. PHƯƠNG PHÁP CHỌN LỰA CÁC YẾU TỐ ẢNH HƯỞNG

3.1. Lựa chọn các yếu tố đầu vào

Yêu cầu đối với các biến được lựa chọn là các yếu tố đầu vào của nghiên cứu thực nghiệm :

- Là các biến độc lập, điều chỉnh được, sự thay đổi giá trị của chúng theo các mức quy hoạch là hoàn toàn độc lập, không phụ thuộc và kéo theo sự thay đổi của các yếu tố khác. Các véc tơ của chúng trong ma trận kế hoạch phải độc lập tuyến tính.

- Là các yếu tố định lượng, vì vậy các yếu tố định tính không có trị số xác định cụ thể như : phương pháp tạo mẫu, màu sắc của đối tượng, hình dạng của bộ phận làm việc ... không thể đưa vào làm yếu tố nghiên cứu của quy hoạch thực nghiệm.

- Có hiệu ứng ảnh hưởng rõ nét đến hàm mục tiêu đánh giá hành vi đối tượng nghiên cứu.

Các căn cứ lựa chọn các yếu tố đầu vào : thông tin tiên nghiệm, kết quả nghiên cứu lý thuyết, ý kiến chuyên gia, các thực nghiệm thăm dò và thực nghiệm sàng lọc.

3.1.1. Thông tin tiên nghiệm

Thông tin có được nhờ kết quả quan sát trực tiếp làm việc của đối tượng nghiên cứu và kết quả tìm hiểu tài liệu tham khảo. Phần lớn các đối tượng nghiên cứu đã được nghiên cứu bằng lý thuyết hoặc thực nghiệm. Đó là những quá trình tương tự diễn ra trong môi trường khác, nhưng có cùng bản chất vật lý, cùng quy luật tác động... Đây là những thông tin sơ bộ, định hướng.

3.1.2. Kết quả nghiên cứu lý thuyết

Trong nhiều trường hợp, người nghiên cứu tuy chưa thể hiểu biết và xây dựng những mô hình lý thuyết cơ bản và toàn diện về đối tượng, nhưng từ những lý thuyết của khoa học cơ sở, hoặc từ các công trình lý thuyết tương tự, có thể mô tả bằng công thức giải tích một số tính chất hoặc hành vi nào đó của đối tượng nghiên cứu.

3.1.3. Ý kiến chuyên gia

Thông thường, thông tin từ các tài liệu rất ít và không toàn diện về đối tượng nghiên cứu. Do vậy có thể sử dụng phương pháp xin ý kiến chuyên gia để đánh giá mức độ quan trọng của các yếu tố ảnh hưởng. Phương pháp này có hiệu quả tốt nếu số yếu tố cần đánh giá lớn và số chuyên gia đông. Đây là phương pháp đã được chuẩn hóa, có thể áp dụng cho nhiều đối tượng nghiên cứu khoa học khác nhau.

3.1.4. Các thực nghiệm thăm dò, thực nghiệm sàng lọc

Đôi khi, sau các bước nói trên vẫn còn lại vài yếu tố ảnh hưởng đáng nghi ngờ mà việc loại bỏ hay giữ lại làm yếu tố nghiên cứu cần nhờ đến kết quả kiểm chứng thực nghiệm. Hoặc khi đối tượng nghiên cứu quá mới mẻ, thông tin ban đầu ít và chưa đủ tin cậy, việc sàng lọc các yếu tố cần tiến hành hết sức cẩn thận. Nếu bỏ sót yếu tố quan trọng xj nào đó, thì các kết quả nghiên cứu sẽ chỉ là 1 thiết diện của mặt mục tiêu tạo bởi mặt phẳng $x_j = \text{const}$. Nhưng trường hợp này đòi hỏi phải tiến hành các thực nghiệm thăm dò.

a. Thực nghiệm thăm dò đơn yếu tố

- Thực hiện thí nghiệm với một yếu tố thay đổi, các yếu tố còn lại được ấn định ở các giá trị cố định.

- Xử lý số liệu trong đó có kiểm tra giả thiết về tính đồng nhất phương sai và đánh giá mức độ ảnh hưởng của yếu tố theo kết quả phân tích phương sai.

- Xác định mô hình toán thực nghiệm đơn yếu tố để tiến hành các phân tích và dự báo cần thiết. Bước này thực hiện theo phương pháp bình phương bé nhất.

b. Thực nghiệm sàng lọc đa yếu tố

Thực nghiệm sàng lọc đa yếu tố cần đáp ứng các yêu cầu:

- Số thí nghiệm so với số yếu tố cần khảo sát là tối thiểu, cho phép đưa vào kế hoạch tối đa các yếu tố thay đổi, mà số thí nghiệm là chấp nhận được, tốn ít công sức.

- Cho phép phân tích và so sánh đối chứng hiệu ứng tác động của các yếu tố riêng rẽ, hoặc các cặp yếu tố theo điều kiện đặt ra ban đầu.

3.2. Phương pháp chuyên gia

Các chuyên gia thuộc nhiều trường phái khác nhau sẽ được đề nghị sắp xếp các yếu tố ảnh hưởng đến đối tượng theo trình tự giảm dần về mức độ ảnh hưởng đến các mục tiêu tối ưu. Mỗi chuyên gia khi được hỏi phải điền vào phiếu điều tra, ở đó đã ghi sẵn các yếu tố, thứ nguyên và khoảng biến thiên của chúng. Các chuyên gia cần phải ghi vị trí thứ tự của mỗi yếu tố càng quan trọng càng có thứ tự hạng trọng số càng lớn, nếu cần thiết có thể bổ sung vào phiếu những yếu tố mới hoặc bỏ bớt yếu tố cũ hoặc nêu ý kiến về miền biến thiên của chúng. Để đảm bảo đánh giá khách quan thì số chuyên gia được hỏi càng nhiều càng tốt.

3.3. Các thực nghiệm sàng lọc theo phương án bão hòa

Sau tất cả các bước: nghiên cứu tài liệu tham khảo, lấy ý kiến chuyên gia, phân tích lý thuyết... nếu số yếu tố ảnh hưởng còn lại khá lớn thì thực nghiệm sàng lọc đóng vai trò sàng lọc quyết định. Tùy theo giả thiết ban đầu, người ta phân thành phương án bão hòa, siêu bão hòa (cân đối ngẫu nhiên) và loại kế tiếp.

Thực nghiệm được gọi là bão hòa khi toàn bộ số bậc tự do của thực nghiệm được dùng để ước lượng các hệ số của mô hình toán thực nghiệm. Giả sử số hệ số trong phương trình hồi quy thực nghiệm là L , số thí nghiệm của thực nghiệm là N , thì thực nghiệm bão hòa là thực nghiệm mà:

$$L = N$$

3.4. Nhóm các yếu tố vào và chọn mục tiêu đánh giá

Mặc dù đã qua các bước sàng lọc, nhưng ở nhiều thực nghiệm số yếu tố cần nghiên cứu còn lại khá lớn. Xét theo quan điểm hệ thống, đưa càng nhiều yếu tố có

ảnh hưởng thực sự vào một kế hoạch thực nghiệm, người nghiên cứu càng có điều kiện tìm được tối ưu có chất lượng cao của đối tượng. Tuy nhiên lại có một số mặt nhược điểm.

Trước hết, khác với thực nghiệm sàng lọc, các thực nghiệm tìm tối ưu ở giai đoạn sau phải đáp ứng các tiêu chuẩn tối ưu nghiêm ngặt của kế hoạch thực nghiệm. Vì thế, khi số yếu tố vào khá lớn (chỉ cần khi $k \geq 7$) thì số thí nghiệm trong kế hoạch đã tăng lên rất nhiều. Ở mỗi điểm lại phải tiến hành một số thí nghiệm song song (lặp lại). Toàn bộ các thí nghiệm cần tiến hành theo trình tự ngẫu nhiên hóa. Các yêu cầu này càng làm tăng khối lượng và thời gian thực nghiệm.

Người nghiên cứu đứng trước sự lựa chọn: xây dựng và tiến hành một kế hoạch thực nghiệm với toàn bộ số yếu tố ảnh hưởng đã chọn, ví dụ $k = 7$, hay tách ra thành 2 kế hoạch song song với $k_1 = 3$; $k_2 = 4$.

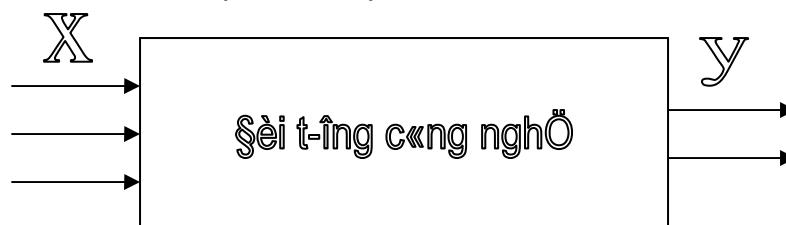
3.5. Ảnh hưởng của các tiên đề của phân tích hồi qui đến sự lựa chọn các yếu tố độc lập

Phân tích hồi qui được xây dựng với những tiên đề mà chúng có liên quan đến môi trường và điều kiện thực nghiệm. Điều kiện và môi trường thực nghiệm lại bị ràng buộc bởi đã điếm các thông số nghiên cứu và chỉ tiêu đánh giá. Mức độ thỏa mãn các tiên đề vcuar phân tích hồi qui phụ thuộc nhiều vào cách chọn nhóm và xác định mức, khoảng biến thiên của các yếu tố ảnh hưởng, vào độ nhạy và độ chính xác của các giá trị quan sát chỉ tiêu đầu ra. Do vậy, người nghiên cứu cần biết rõ các yêu cầu này để có những quyết định tối ưu ngay ở bước xác định các yếu tố nghiên cứu.

Chương 4. CÁC PHƯƠNG PHÁP HOẠCH ĐỊNH THỰC NGHIỆM

4.1. QUY HOẠCH TRỰC GIAO CẤP I

CÁC BƯỚC QUY HOẠCH TRỰC GIAO CẤP I



1. XÁC ĐỊNH MIỀN BIẾN THIÊN

$$Z_j^{\min} < Z_j < Z_j^{\max}$$

và **TÂM QUY HOẠCH** : $Z_j^o = 0.5(Z_j^{\min} + Z_j^{\max})$

2. CHỌN DẠNG PHƯƠNG TRÌNH HỒI QUY

sau khi đã mã hóa : $x_j = 2(Z_j - Z_j^o) / (Z_j^{\max} - Z_j^{\min})$

+ chọn dạng tuyến tính :

$$y_1 = b_0 + b_1x_1 + \dots + b_kx_k$$

(+ hoặc dạng :

$$y_1 = b_0 + b_1x_1 + \dots + b_kx_k + b_{12}x_1x_2 + \dots + b_{k-1,k}x_{k-1}x_k)$$

3. THỰC HIỆN N THÍ NGHIỆM $N = 2^k$

TÍNH TOÁN XÁC ĐỊNH CÁC HỆ SỐ HỒI QUY b_j

bằng phương pháp **Bình phương cực tiểu**

4. KIỂM ĐỊNH SỰ CÓ NGHĨA CỦA CÁC HỆ SỐ HỒI QUY b_j với chuẩn Student

Thực hiện các thí nghiệm tại tâm quy hoạch hoặc sử dụng các thí nghiệm song song, lặp lại. Loại bỏ các b_j không có nghĩa, tính toán lại các b_j và kiểm định lại cho tới khi chỉ còn các b_j có nghĩa

5. KIỂM ĐỊNH SỰ CÓ NGHĨA CỦA PHƯƠNG TRÌNH HỒI QUY với chuẩn Fisher

4.1.1. Thực nghiệm yếu tố toàn phần TYT 2^k

Tuỳ thuộc thông tin ban đầu mà người nghiên cứu tổ chức các thí nghiệm để nhận được mô hình thống kê thực nghiệm dạng tuyến tính hoặc phi tuyến

Chọn qui hoạch thực nghiệm yếu tố toàn phần và từng phần

Những thực nghiệm mà mọi tổ hợp các mức của yếu tố đều được thực hiện để nghiên cứu gọi là thực nghiệm yếu tố toàn phần (TYT n^k)

$$N = n^k \quad (4.1)$$

Trong đó : N : lượng thí nghiệm

n : số lượng mức của các yếu tố

k : số yếu tố ảnh hưởng

Xét thực nghiệm yếu tố toàn phần 2 mức k yếu tố ảnh hưởng

4.1.1.1. Cách tổ chức thí nghiệm trực nghiệm trực giao cấp I

1 – Số thí nghiệm cần thực hiện

$$N = 2^k$$

2 – Mức cơ bản

$$Z_j^0 = \frac{Z_j^{\max} + Z_j^{\min}}{2} \quad (4.2)$$

Trong đó: Z_j^0 là mức cơ bản (tâm phương án).

Z_j^{\max} là mức trên (mức cao).

Z_j^{\min} là mức dưới (mức thấp).

Vectơ vào tại mức cơ bản Z_j^0 ($j = 1, 2, \dots, k$) chỉ ra không gian các yếu tố của một điểm đặc biệt gọi là tâm thực nghiệm.

3 – Khoảng biến thiên

$$\lambda_j = \frac{Z_j^{\max} - Z_j^{\min}}{2} \quad (j = 1, 2, 3 \dots k) \quad (4.3)$$

λ_j là khoảng biến thiên theo trục Z_j .

Ví dụ: Xem trang 25 (Giáo trình “Quy hoạch thực nghiệm nghiên cứu và ứng dụng” của Nguyễn Thị Lan).

4 – Biến không thứ nguyên : kí hiệu x_j

Mã hoá được thực hiện dễ dàng nhờ việc chọn tâm Z_j^0 của miền nghiên cứu làm gốc toạ hệ trục độ.

$$\begin{cases} x_j^{\max} = \frac{Z_j^{\max} - Z_j^0}{\lambda_j} \\ x_j^{\min} = \frac{Z_j^{\min} - Z_j^0}{\lambda_j} \\ x_j^0 = \frac{Z_j^0 - Z_j^0}{\lambda_j} \end{cases} \quad j = 1, 2, 3, \dots, k \quad (4.4)$$

⇒ Tọa độ thứ nguyên mức trên (x_j^{\max}) luôn bằng +1; mức dưới (x_j^{\min}) luôn bằng -1 và tọa độ của tâm phương án (x_j) luôn bằng 0 và trùng với gốc tọa độ.

$$\begin{cases} Z_j^{\max} = Z_j^0 + \lambda_j \\ Z_j^{\min} = Z_j^0 - \lambda_j \end{cases} \quad (4.5)$$

5 – Lập ma trận thực nghiệm

Ma trận thực nghiệm với biến thực nghiệm là một dạng mô tả chuẩn các điều kiện tiến hành thí nghiệm theo bảng chữ nhật. Mỗi hàng là một thí nghiệm, trong ma trận có một số hàng giống nhau mà thông số đều ở mức cơ sở Z_j^0

Ma trận thực nghiệm với biến ảo là ma trận chỉ bao gồm các biến ảo x_j . Khi xây dựng ma trận thực nghiệm đưa thêm biến $x_0 = \pm 1$ và bố trí các thí nghiệm sao cho không có thí nghiệm nào trùng nhau. Theo kinh nghiệm làm như sau :

- Xác định số thí nghiệm cần thực hiện theo công thức $N = 2^k$, cột x_0 luôn bằng 1

- Lập cho từng yếu tố ảnh hưởng và lần lượt từ x_1 đến x_k .

Chú ý : Người nghiên cứu nên đưa các thí nghiệm ở tâm vào ma trận

Tính Y_0 (giá trị ở tâm thực nghiệm); $b_0 \rightarrow$ có thể dự đoán được vùng nghiên cứu thuộc vùng tuyến tính hoặc phi tuyến.

6 – Tính chất ma trận trực giao cấp I

Ma trận trực giao cấp I có những tính chất sau:

- Tính đối xứng qua tâm thực nghiệm.

$$\sum_{u=1}^N x_{iu} = 0 \quad ; \quad i = 1, 2, 3, \dots, k \quad (4.6)$$

$$u = 1, 2, 3, \dots, N$$

- Tính trực giao giữa 2 cột trong ma trận thực nghiệm.

$$\sum_{u=1}^N x_{iu}x_{ju} = 0 \quad ; \quad i \neq j = 1,2,3\dots k \quad (4.7)$$

- Tính bất biến khi quay hệ trục quanh tâm thực nghiệm.

$$\sum_{u=1}^N x_{iu}^2 = N \quad ; \quad i = 1,2,3,\dots k \quad (4.8)$$

* Ưu điểm của ma trận trực giao cấp I:

- Khi loại bỏ những hệ số không có nghĩa sẽ không phải tính lại các hệ số có nghĩa.

- Phương sai các hệ số b ($S_{b_j}^2$) trong phương trình hồi qui có giá trị tối thiểu, xác định theo kết quả của N thí nghiệm và nhỏ hơn phương sai tái hiện S_{th}^2 .

- Tâm phương án thông tin nhiều nhất \rightarrow chỉ lần thực nghiệm lặp ở tâm thực nghiệm là đủ.

4.1.1.2. Một số dạng của phương trình hồi qui cấp I

Trước tiên, phải biết được sự phụ thuộc giữa các thông số đầu vào và thông số đầu ra $Y=f(x)$ để chọn phương trình hồi qui sao cho hợp lý.

Đối với qui hoạch thực nghiệm, những phương trình hồi qui cấp I thường chọn các khai triển của đa thức có dạng tổng quát sau :

$$\tilde{Y} = b_0 + b_1x_1 + \dots + b_kx_k + \dots + b_{ij}x_ix_j + \dots + b_{ijk}x_ix_jx_k \quad ; \quad \text{với } i \neq j \neq k = 1,2,3\dots k \quad (4.9)$$

Trong đó: b_0 là hệ số hồi qui.

b_j là hệ số tuyến tính.

$b_{ij} ; b_{ijk}$ là hệ số tương tác cặp và tương tác ba.

Để đơn giản thì chọn dạng phương trình hồi qui dạng tuyến tính.

Muốn xây dựng phương trình hồi qui đầy đủ đưa thêm vào phương trình tuyến tính các hệ số tương tác

Với $k = 2$ (2 yếu tố ảnh hưởng) ta có:

$$\begin{cases} \tilde{Y} = b_0 + b_1x_1 + b_2x_2 \\ \tilde{Y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 \end{cases} \quad (4.10)$$

Với $k = 3$ ta có:

$$\begin{cases} \tilde{Y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \\ \tilde{Y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{23}x_1x_2 + b_{13}x_1x_3 + b_{123}x_1x_2x_3 \end{cases} \quad (4.11)$$

4.1.1.3. Lập công thức tính hệ số b trong phương trình hồi qui

1 – Phương án bình phương nhỏ nhất (BPNN)

Là phương án cơ bản có hiệu lực khi xử lí các số liệu thực nghiệm và xây dựng mô hình thống kê cho nhiều đối tượng nghiên cứu thuộc các lĩnh vực khác nhau.

Phương pháp này cho phép xác định các hệ số của phương trình hồi qui đã chọn sao cho độ lệch của sự phụ thuộc đã cho so với số liệu thực nghiệm là nhỏ nhất.

$$\Phi = \sum_{u=1}^N (\tilde{Y}_u - Y_u)^2 \rightarrow \min \quad (4.12)$$

Trong đó : Y_u là giá trị thực nghiệm ứng với k thông số tối ưu ở thí nghiệm thứ u .

\tilde{Y}_u là giá trị theo phương trình hồi qui số tối ưu ở thí nghiệm thứ u .

2 – Hệ phương trình chuẩn tắc

Xét $k = 2$, dạng PTHQ như sau:

$$\tilde{Y} = b_0x_{0u} + b_1x_{1u} + b_2x_{2u} + b_{12}x_{1u}x_{2u} \quad (4.13)$$

Thay (4.13) vào (4.12):

$$\Phi = \sum_{u=1}^N [(b_0x_{0u} + b_1x_{1u} + b_2x_{2u} + b_{12}x_{1u}x_{2u}) - Y_u]^2 \rightarrow \min \quad (4.14)$$

Φ cực tiểu khi thỏa mãn các điều kiện sau:

$$\frac{\partial \Phi}{\partial b_0} = 0 \quad ; \quad \frac{\partial \Phi}{\partial b_1} = 0 \quad ; \quad \frac{\partial \Phi}{\partial b_2} = 0 \quad ; \quad \frac{\partial \Phi}{\partial b_{12}} = 0 \quad (4.15)$$

Có thể viết dưới dạng sau:

$$\begin{cases} \frac{\partial \Phi}{\partial b_0} = \sum_{u=1}^4 [(b_0 x_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2)u - Y_u] x_{0u} = 0 \\ \frac{\partial \Phi}{\partial b_1} = \sum_{u=1}^4 [(b_0 x_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2)u - Y_u] x_{1u} = 0 \\ \frac{\partial \Phi}{\partial b_2} = \sum_{u=1}^4 [(b_0 x_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2)u - Y_u] x_{2u} = 0 \\ \frac{\partial \Phi}{\partial b_{12}} = \sum_{u=1}^4 [(b_0 x_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2)u - Y_u] x_{1u} x_{2u} = 0 \end{cases} \quad (4.16)$$

$$\begin{cases} b_0 \sum_{u=1}^4 x_{0u} + b_1 \sum_{u=1}^4 x_{1u} + b_2 \sum_{u=1}^4 x_{2u} + b_{12} \sum_{u=1}^4 x_{1u} x_{2u} = \sum_{u=1}^4 Y_u x_{0u} \\ b_0 \sum_{u=1}^4 x_{1u} + b_1 \sum_{u=1}^4 x_{1u}^2 + b_2 \sum_{u=1}^4 x_{1u} x_{2u} + b_{12} \sum_{u=1}^4 x_{1u}^2 x_{2u} = \sum_{u=1}^4 Y_u x_{1u} \\ b_0 \sum_{u=1}^4 x_{2u} + b_1 \sum_{u=1}^4 x_{1u} x_{2u} + b_2 \sum_{u=1}^4 x_{2u}^2 + b_{12} \sum_{u=1}^4 x_{1u} x_{2u}^2 = \sum_{u=1}^4 Y_u x_{2u} \\ b_0 \sum_{u=1}^4 x_{1u} x_{2u} + b_1 \sum_{u=1}^4 x_{1u}^2 x_{2u} + b_2 \sum_{u=1}^4 x_{1u} x_{2u}^2 + b_{12} \sum_{u=1}^4 (x_{1u} x_{2u})^2 = \sum_{u=1}^4 Y_u x_{1u} x_{2u} \end{cases} \quad (4.17)$$

Phương trình (4.17) gọi là hệ phương trình chuẩn tắc.

3 – Công thức tính hệ số b của phương trình hồi qui

Các hệ số b trong phương trình hồi qui độc lập nhau và xác định như sau:

$$\begin{cases} b_0 = \frac{1}{4} \sum_{u=1}^4 x_{0u} Y_u \\ b_1 = \frac{1}{4} \sum_{u=1}^4 x_{1u} Y_u \\ b_2 = \frac{1}{4} \sum_{u=1}^4 x_{2u} Y_u \\ b_{12} = \frac{1}{4} \sum_{u=1}^4 x_{1u} x_{2u} Y_u \end{cases} \quad (4.18)$$

⇒ Công thức tổng quát để tính các hệ số b trong phương trình hồi qui của qui hoạch trực giao cấp I tương ứng với k yếu tố ảnh hưởng như sau:

$$\left\{ \begin{array}{l} b_0 = \frac{1}{N} \sum_{u=1}^4 x_{0u} Y_u \\ b_j = \frac{1}{N} \sum_{u=1}^4 x_{ju} Y_u \\ b_{ij} = \frac{1}{N} \sum_{u=1}^4 x_{iu} x_{ju} Y_u \\ b_{ijl} = \frac{1}{N} \sum_{u=1}^4 x_{iu} x_{ju} x_{lu} Y_u \end{array} \right. \quad i \neq j \neq l = 1, 2, 3 \dots k \quad (4.19)$$

4 – Ý nghĩa của hệ số b trong phương trình hồi qui

Giá trị của hệ số b_j trong phương trình hồi qui đặc trưng cho sự đóng góp của yếu tố thứ j vào đại lượng Y.

Hệ số nào có giá trị tuyệt đối lớn nhất thì yếu tố tương ứng sẽ ảnh hưởng đến quá trình là nhiều nhất.

4.1.1.4. Kiểm tra ý nghĩa của các hệ b trong phương trình hồi qui

Để kiểm tra ý nghĩa của các hệ số b trong phương trình hồi qui phải tính phương sai tái hiện (làm thí nghiệm song song ở mỗi điểm thực nghiệm).

Hệ số b trong phương trình hồi qui độc lập nhau và xác định với một độ chính xác (S_{bj}).

$$S_{bj} = \frac{S_{th}}{\sqrt{N}} \quad (4.20)$$

N: số thí nghiệm ứng mỗi phương án.

Tính ý nghĩa của các hệ số b được kiểm định theo chuẩn Student (t)

xác định như sau :

$$t_j = \frac{|b_j|}{S_{bj}} \quad (4.21)$$

Trong đó: b_j là hệ số thứ j trong phương trình hồi qui tính theo (4.19).

S_{bj} : độ lệch quân phương của hệ số j được xác định theo công thức (4.20)

Các bước kiểm tra được tiến hành như mục kiểm định thống kê (chương 2)

Công thức (4.21) xác định được S_{bj} ứng với mỗi phương án thực nghiệm.

1 – Phương án thực nghiệm tại tâm

Khi hoàn tất 2^k thí nghiệm ở nhân phương án, người nghiên cứu phải làm thêm m (m ít nhất bằng 3) thí nghiệm ở tâm phương án với các giá trị ứng với thí nghiệm tâm là: $Y_1^0, Y_2^0, Y_3^0 \dots$

Phương sai tái hiện được xác định:

$$S_{th}^2 = \frac{\sum_{i=1}^m (Y_i^0 - \bar{Y}^0)^2}{m-1} \quad i = 1, 2, 3 \dots m \quad (4.22)$$

$$S_{th} = \sqrt{S_{th}^2} \quad (4.23)$$

Trong đó : Y_i^0 là giá trị đo được ở lần lặp thứ i
 \bar{Y}^0 là giá trị trung bình của m lần đo
 m : số lần lặp

Thay (4.23) vào (4.20) tìm được giá trị S_{bj} .

2 – Phương án thí nghiệm song song

Tại mỗi điểm thí nghiệm được lặp lại m lần. Trước khi tính toán hệ số b và kiểm định các thông số thông kê phải kiểm tra sự đồng nhất của các phương sai theo chuẩn Cochran (G), chỉ được phép ước lượng các sai số khi phương sai đồng nhất.

Phương sai tái hiện của một cuộc thí nghiệm:

$$S_{th}^2 = \frac{\sum_{u=1}^N \sum_{i=1}^m (Y_{ui} - \bar{Y}_i)^2}{N(m-1)} \quad (4.24)$$

Phương sai phân phối trung bình của một cuộc thí nghiệm

$$S_{th}^2(\bar{Y}) = \frac{S_{th}^2}{m} \quad (4.25)$$

Phương sai của hệ số b_j

$$S_{bj}^2 = \frac{S_{th}^2(\bar{Y})}{N} \quad (4.26)$$

Sai số chuẩn (độ lệch quân phương) của hệ số b_j

$$S_{bj} = \frac{S_{th}(\bar{Y})}{\sqrt{N}} \quad (4.27)$$

Sau khi kiểm tra ý nghĩa của các hệ số b_j , viết PTHQ với các hệ số có nghĩa và kiểm tra tương thích của phương trình hồi qui với thực nghiệm.

4.1.1.5. Kiểm tra sự tương thích của PTHQ với thực nghiệm

Sự tương thích của PTHQ với thực nghiệm được kiểm định theo chuẩn Fisher (E). Các bước kiểm tra được trình bày ở mục kiểm định thống kê (chương 2).

$$F = \frac{S_{tt}^2}{S_{th}^2} \quad (4.30)$$

Đối với phương án thí nghiệm tại tâm

$$S_{tt}^2 = \frac{\sum_{u=1}^N (Y_u - \tilde{Y}_u)^2}{f_{tt}} \quad (4.31)$$

Phương án thí nghiệm song song, với lần lặp ở mỗi điểm thực nghiệm là m

$$S_{tt}^2 = \frac{m \sum_{u=1}^N (Y_u - \tilde{Y}_u)^2}{f_{tt}} \quad (4.32)$$

Y_u, \bar{Y}_u : là giá trị thực nghiệm.

Y_u : giá trị tính theo PTHQ.

f_{tt} : độ tự do ứng với phương sai tương thích (S_{tt}^2).

$$f_{tt} = N - L$$

N : số thí nghiệm trong phương án.

L : số hệ số có nghĩa được kiểm tra ở mục (4.1.1.4).

Sau khi kiểm tra nếu PTHQ tương thích với thực nghiệm sẽ được sử dụng để tìm kiếm tối ưu. Nếu không phù hợp sẽ phải xem xét lại từng bước của bài qui hoạch và chọn mô tả toán học ở mức cao hơn.

4.1.2. Thực nghiệm yếu tố từng phần 2^{k-p}

Để mô tả quá trình thực nghiệm thì qui hoạch thực nghiệm yếu tố toàn phần 2^k không hiệu quả khi số yếu tố k khá lớn. Số yếu tố k tăng chậm mà số thí nghiệm tăng quá nhanh ($N=2^k$) và sẽ có rất nhiều bậc tự do để kiểm tra sự tương thích của PTHQ với thực nghiệm.

Thật khó khăn về kinh tế khi phải thực hiện 1 cuộc thí nghiệm 2^k mà yếu tố $k > 4$.

Vì vậy sẽ giảm đáng kể số thí nghiệm nếu ta dùng thực nghiệm yếu tố từng phần (lời giải từng phần) mà người nghiên cứu vẫn thu được mô hình thí nghiệm mô tả tương thích quá trình thí nghiệm.

Kí hiệu: TYP 2^{k-p}

Trong đó:

2: là 2 mức của mỗi yếu tố ảnh hưởng

k: số yếu tố ảnh hưởng

p: đặc trưng cho mức độ từng phần

4.1.2.1. Xây dựng mô hình thống kê thực nghiệm

4.1.2.2 Cách tổ chức thí nghiệm trong phương án thực nghiệm từng phần

$$N=2^{k-p}$$

Số thí nghiệm trong phương án từng phần bằng $\frac{1}{2^p}$ bảng TYT 2^k

4.1.2.3. Công thức tính hệ số b trong PTHQ của qui hoạch phân bảng TYP 2^{k-p}

Để cho lời giải từng phần là một phương án trực giao ta cần chọn phương án thực nghiệm yếu tố toàn phần có số yếu tố ảnh hưởng nhỏ hơn làm mức cơ sở.

Được áp dụng giống như trong qui hoạch TYT 2^k

$$b_0 = \frac{1}{N} \sum_{u=1}^N Y_u X_{0u}$$

$$b_j = \frac{1}{N} \sum_{u=1}^N Y_u X_{ju}$$

$$u = 1, 2, \dots, N$$

$$j = 1, 2, \dots, k$$

4.1.2.4. Các bước thực hiện qui hoạch phân bảng

1/ Trường hợp k=3, p=1

- Lập qui hoạch và xây dựng ma trận TYT 2^2 .
- Thay cột có hiệu ứng tương tác bằng hiệu ứng tuyến tính ($x_1 x_2 = x_3$).
- Làm 4 thí nghiệm và dùng kết quả của 4 thí nghiệm để tính hệ số b_0, b_1, b_2, b_3 .

- Sau khi làm 4 thí nghiệm đầu, vì một lý do nào đó người nghiên cứu cho rằng tương tác cặp có ý nghĩa thì làm 4 thí nghiệm của nửa bảng còn lại, nhưng lần này thay yếu tố bổ sung $x_3 = -x_1x_2$. Như vậy qui hoạch sẽ còn nửa bảng.

$$x_3 = x_1x_2$$

$$x_3 = -x_1x_2$$

2/ Trường hợp $k=4, p=1$

- Lập qui hoạch TYT 2^3

- Thay $x_3 = x_1x_2x_3$

- Làm 8 thí nghiệm và sử dụng kết quả của 8 thí nghiệm để tính hệ số b_0, b_1, b_2, b_3, b_4 .

Như vậy qui hoạch phân bảng với 2 nửa bảng khi thay $x_4 = x_1x_2x_3$

3/ Trường hợp $k=5, p=2$

Lập qui hoạch TYT 2^3 .

- Thay $x_4 = x_1x_2$, (bỏ qua tương tác x_1x_2), $x_5 = x_1x_2x_3$ (bỏ qua tương tác $x_1x_2x_3$).

- Làm 8 thí nghiệm và dùng kết quả của thí nghiệm đó để xác định hệ số b_0 và 5 hệ số còn lại.

- Qui hoạch phân bảng với 4 phần bảng như sau:

$$1 \{x_4 = x_1x_2, x_5 = x_1x_2x_3\}$$

$$2 \{x_4 = -x_1x_2, x_5 = -x_1x_2x_3\}$$

$$3 \{x_4 = x_1x_2, x_5 = -x_1x_2x_3\}$$

$$4 \{x_4 = -x_1x_2, x_5 = x_1x_2x_3\}$$

4.2. QUY HOẠCH TRỰC GIAO CẤP II

CÁC BƯỚC QUY HOẠCH TRỰC GIAO CẤP II



6. **XÁC ĐỊNH MIỀN BIẾN THIÊN** $Z_j^{\min} < Z_j < Z_j^{\max}$

và **TÂM QUY HOẠCH**: $Z_j^0 = 0.5(Z_j^{\min} + Z_j^{\max})$

7. **CHỌN DẠNG PHƯƠNG TRÌNH HỒI QUY**

sau khi đã mã hóa: $x_j = 2(Z_j - Z_j^0) / (Z_j^{\max} - Z_j^{\min})$

$$y_1 = b_0 + b_1x_1 + \dots + b_kx_k + b_{12}x_1x_2 + \dots + b_{11}x_1^2 + \dots + b_{kk}x_k^2$$

8. **THỰC HIỆN N THÍ NGHIỆM** $N = 2^k + 2k + n_0$

Trong đó: - 2^k thí nghiệm của QHTG cấp I

với các $Z_j = Z_j^{\min}$ hoặc $Z_j = Z_j^{\max}$

- $2k$ thí nghiệm tại các điểm "sao": $x_j = \text{TG}$

hoặc $x_j = -\text{TG}$

- n_0 thí nghiệm tại tâm $Z_j = Z_j^0$

9. **TÍNH TOÁN XÁC ĐỊNH CÁC HỆ SỐ HỒI QUY b_j**

bằng phương pháp **Bình phương cực tiểu**

10. **KIỂM ĐỊNH SỰ CÓ NGHĨA CỦA CÁC HỆ SỐ HỒI QUY b_j**

với **chuẩn Student**

Thực hiện các thí nghiệm tại tâm quy hoạch hoặc sử dụng các thí nghiệm song song. Loại bỏ các b_j không có nghĩa, tính toán lại các b_j và kiểm định lại cho tới khi chỉ còn các b_j có nghĩa

11. **KIỂM ĐỊNH SỰ CÓ NGHĨA CỦA PHƯƠNG TRÌNH HỒI QUY** với **chuẩn Fisher**

4.3. Tối ưu hóa qui hoạch thực nghiệm

Bước 1

- Xác định một điểm xuất phát nằm trong miền giới hạn tổng thể của các biến đầu vào. Chọn điểm đó làm mức cơ bản, chọn khoảng biến thiên của từng biến để xác định miền giới hạn của quy hoạch thực nghiệm trực giao cấp một.

Bước 2

- Làm các thí nghiệm theo quy hoạch trực giao cấp một
- Xây dựng phương trình hồi quy bậc nhất .

Nếu phương trình hồi quy bậc nhất **không tương thích** thì chuyển tới thực hiện bước 4 .

Nếu phương trình hồi quy bậc nhất **tương thích** thì thực hiện bước 3.

Bước 3

- Xác định vectơ **gradient** của hàm mục tiêu tại mức cơ bản và xuất phát từ mức cơ bản xác định tọa độ các điểm thực nghiệm nằm cách đều nhau trên hướng của vectơ gradient với khoảng cách tự chọn phù hợp với đối tượng nghiên cứu. Làm thực nghiệm để xác định một điểm có giá trị hàm mục tiêu tốt nhất trên hướng gradient. Chọn điểm tìm được làm điểm xuất phát mới và quay về bước 2 .

Bước 4

- Làm các thí nghiệm theo **quy hoạch cấp hai** (trực giao hoặc quay).

Bước 5

- Xây dựng phương trình hồi quy bậc hai.

- Nếu phương trình hồi quy bậc hai **không tương thích** thì chuyển tới thực hiện bước 6 .

- Nếu phương trình hồi quy bậc hai **tương thích** thì thực hiện bước 7.

Bước 6

- Thu hẹp khoảng biến thiên của các biến đầu vào rồi quay về bước 5.

Bước 7

- **Tìm cực trị của hàm mục tiêu thu được** ở dạng phương trình hồi quy bậc hai thu được ở **bước 5** và làm lại thực nghiệm để kiểm chứng và đánh giá kết quả.

4.3.1. Tối ưu hóa theo phương pháp leo dốc

Bước 1: Chọn điểm xuất phát $X^{(0)} (x_1^{(0)}, \dots, x_n^{(0)})$

Chọn các giá trị $\varepsilon_y > 0$ và $\varepsilon_x > 0$

Xác định $y(X^{(0)})$

Bước 2: Xác định vector gradient tại điểm $X^{(0)}$

Bước 3 : Chọn số λ dương;

Từ điểm $X^{(0)}$ xác định $X^{(1)}$:

$$x_1^{(1)} = x_1^{(0)} \pm \lambda \left. \frac{\partial y}{\partial x_1} \right|_{X=X^0}$$

$$x_2^{(1)} = x_2^{(0)} \pm \lambda \left. \frac{\partial y}{\partial x_2} \right|_{X=X^0}$$

.....

$$x_n^{(1)} = x_n^{(0)} \pm \lambda \left. \frac{\partial y}{\partial x_n} \right|_{X=X^0}$$

(dấu “+” khi tìm max , dấu “-” khi tìm min)

Xác định $y(X^{(1)})$

Bước 4: So sánh $y(X^{(1)})$ với $y(X^{(0)})$.

Nếu $y(X^{(1)})$ ‘tốt’ hơn $y(X^{(0)})$ \rightarrow tiếp tục lặp lại bước 3 để leo dốc tới $X^{(2)}$, $X^{(3)}$, ..., $X^{(k)}$

Nếu $y(X^{(k)})$ “xấu” hơn $y(X^{(k-1)}) \rightarrow$ Thực hiện phép gán $X^{(1)} = X^{(k-1)}$ và $y^{(1)} = y(X^{(k-1)})$, sau đó chuyển sang **bước 5**

Bước 5: Kiểm tra điều kiện dừng:

$$|y^{(1)} - y^{(0)}| \leq \varepsilon_y \quad \text{hoặc / và} \quad \sqrt{(x_1^{(1)} - x_1^{(0)})^2 + \dots + (x_n^{(1)} - x_n^{(0)})^2} \leq \varepsilon_x \quad (*)$$

- Nếu (*) không thỏa mãn:

+ Chọn $X^{(1)}$ làm điểm xuất phát mới (nói cách khác: thực hiện phép gán $X^{(0)} = X^{(1)}$ và $y^{(0)} = y^{(1)}$)

+ Quay lại **bước 2**

- Nếu (*) thỏa mãn \rightarrow **kết luận: y đạt giá trị tối ưu tại $X^{(1)}$**

4.3.2. Phương pháp luân phiên từng biến giải bài toán tối ưu phỏng định

* **Bước 1:** Chọn điểm xuất phát $X^{(0)} (x_1^{(0)}, \dots, x_n^{(0)})$,

Chọn các giá trị $\varepsilon_y > 0$ và $\varepsilon_x > 0$

Làm thực nghiệm xác định giá trị $y^{(0)}$

* **Bước 2:** Thực hiện n phiên giải bài toán tối ưu lần lượt với từng biến x_i để từ điểm xuất phát $X^{(0)} (x_1^{(0)}, \dots, x_n^{(0)})$ tìm ra điểm $X^{(1)} (x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)})$ tốt hơn.

- Phiên 1: Cố định (n-1) biến, giải bài toán tối ưu với biến còn lại (giả sử x_1) khi cho x_1 chạy trong miền giá trị của nó. Giả sử y tốt nhất tại $X^{(*1)} = (x_1^{(1)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)})$

- Phiên 2: Tiến hành tương tự với biến x_2 (cố định các biến còn lại trong đó $x_1 = x_1^{(1)}$). Tìm được giá trị y tốt nhất tại điểm $X^{(*2)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(0)}, \dots, x_n^{(0)})$.

- Phiên thứ k: Giải bài toán tối ưu với biến x_k (cố định các biến còn lại trong đó $x_1 = x_1^{(1)}, \dots, x_{k-1} = x_{k-1}^{(1)}, x_{k+1} = x_{k+1}^{(0)}, x_n = x_n^{(0)}$). Tìm được giá trị y tốt nhất tại điểm $X^{(*k)} = (x_1^{(1)}, \dots, x_k^{(1)}, x_{k+1}^{(0)}, \dots, x_n^{(0)})$.

- Phiên thứ n : Giải bài toán tối ưu với biến x_n (cố định các biến còn lại trong đó $x_1 = x_1^{(1)}, \dots, x_{k-1} = x_{k-1}^{(1)}, x_{k+1} = x_{k+1}^{(0)}, \dots, x_n = x_n^{(0)}$). Tìm được giá trị y tốt nhất tại điểm $X^{(*n)} = (x_1^{(1)}, \dots, x_k^{(1)}, x_{k+1}^{(1)}, \dots, x_n^{(1)})$.

Đặt $X^{(1)} = X^{(*n)}; y^{(1)} = y(X^{(1)})$

* **Bước 3**: Kiểm tra điều kiện dừng:

$$|y^{(1)} - y^{(0)}| \leq \varepsilon_y \quad \text{hoặc/và} \quad \sqrt{(x_1^{(1)} - x_1^{(0)})^2 + \dots + (x_n^{(1)} - x_n^{(0)})^2} \leq \varepsilon_x \quad (*)$$

trong đó $y^{(1)} = y(X^{(1)}) = y(x_1^{(1)}, \dots, x_n^{(1)})$

- Nếu (*) không thỏa mãn:

+ Chọn $X^{(1)}$ làm điểm xuất phát mới (nói cách khác : thực hiện phép gán $X^{(0)} = X^{(1)}$ và $y^{(0)} = y^{(1)}$)

+ Quay lại bước 2

- Nếu (*) thỏa mãn: kết luận y đạt giá trị tối ưu tại $X^{(1)}$

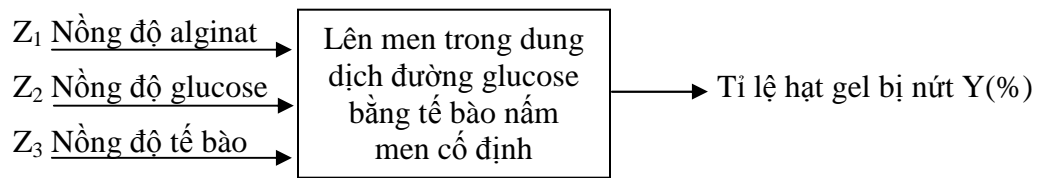
Chương 5. ỨNG DỤNG QUY HOẠCH THỰC NGHIỆM TRONG CÁC QUÁ TRÌNH CÔNG NGHỆ HÓA HỌC

5.1. Bài toán 1.

➤ Mục đích: Nghiên cứu tối ưu hoá quy trình cố định tế bào nấm men bằng Alginat để lên men rượu.

➤ Quy trình công nghệ được mô tả theo sơ đồ (trang 1).

Nghiên cứu các yếu tố ảnh hưởng đến mạng lưới gel: nồng độ alginat; nồng độ glucose; nồng độ tế bào:



Sau quá trình lên men, vớt các hạt gel ra và xác định tỉ lệ (%) hạt gel bị nứt. Tỉ lệ hạt gel bị nứt càng thấp càng tốt nghĩa là hạt gel càng chắc càng tốt.

➤ Hàm mục tiêu: $Y = Y(Z_1, Z_2, Z_3)$

Bài toán tối ưu: Xác định nồng độ alginat; nồng độ glucose; nồng độ tế bào nấm men để hạt gel bền nhất trong quá trình lên men rượu bằng tế bào nấm men, cố định bằng alginat.

$$Y_{\min} = \min Y(Z_1, Z_2, Z_3)$$

Sau khi tiến hành các thí nghiệm thăm dò, tác giả đã chọn vùng khảo sát như sau:

$$Z_1 = 1 \div 4\%$$

$$Z_2 = 10 \div 18\%$$

$$Z_3 = 10 \div 20\%$$

Đây là bài toán tối ưu phỏng định, giải bài toán theo phương pháp leo dốc.

Phương án qui hoạch thực nghiệm: phương pháp trực giao cấp 1.

Số thí nghiệm phải làm: $2^k = 2^3 = 8$

$$\text{Với } Z_{1\min} = 1 \quad Z_1 \quad 4 = Z_{1\max}$$

$$Z_{2\min} = 10 \quad Z_2 \quad 18 = Z_{2\max}$$

$$Z_{3\min} = 10 \quad Z_3 \quad 20 = Z_{3\max}$$

Điểm xuất phát ở tâm phương án:

$$Z_0 = (2.5; 14; 15)$$

Giá trị của hàm mục tiêu tại điểm Z_0 được xác định bằng thực nghiệm:

$$Y(Z_0) = 7.5$$

Ma trận thực nghiệm được bố trí như sau:

Số TN	Z_1	Z_2	Z_3	Y
1	4	18	20	12.35
2	4	18	10	8.87
3	4	10	20	12.08
4	4	10	10	6.92
5	1	18	20	42.13
6	1	18	10	13.51
7	1	10	20	22.19
8	1	10	10	4.57

Phương trình hồi qui có dạng:

$$Y = B_0 + B_1Z_1 + B_2Z_2 + B_3Z_3$$

Trong hệ mã hoá không thứ nguyên ta có được:

Mức trên: - kí hiệu +1

Mức cơ sở: - kí hiệu 0

Mức dưới: - kí hiệu -1

Công thức chuyển từ hệ đơn vị thực qua đơn vị mã hoá không thứ nguyên:

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

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

Thu được ma trận thực nghiệm với các biến mã như sau:

Số TN	X_0	X_1	X_2	X_3	Y
1	1	1	1	1	12.35
2	1	1	1	-1	8.87

3	1	1	-1	1	12.08
4	1	1	-1	-1	6.92
5	1	-1	1	1	42.13
6	1	-1	1	-1	13.51
7	1	-1	-1	1	22.19
8	1	-1	-1	-1	4.57

Từ kết quả thực nghiệm, tính toán các hệ số B_j :

$$B_0 = \frac{\sum_{i=1}^n Y_i}{N}$$

$$B_i = \frac{\sum_{i=1}^n X_{ij} Y_i}{N}$$

$$B_{j1} = \frac{\sum_{i=1}^n (X_j \cdot X_i) \cdot Y_i}{N}$$

Từ số liệu thực nghiệm trên, áp dụng các công thức trên ta xác định được giá trị B_0 , B_1 , B_2 và B_3 thu được kết quả:

B_1	B_2	B_3	B_0
-5.2725	3.8875	6.86	15.3275

➤ Để tính phương sai tái hiện, tác giả làm thêm 3 thí nghiệm ở tâm.

Kết quả các thí nghiệm ở tâm:

N_0	Y_u^0	\bar{Y}^0	$Y_u^0 - \bar{Y}^0$	$(Y_u^0 - \bar{Y}^0)^2$	$\Sigma(Y_u^0 - \bar{Y}^0)^2$
1	5.65	7.50333	-1.8533	3.43472	8.22749
2	7.19		-0.3133	0.09818	
3	9.67		2.1667	4.69459	

Phương sai tái hiện được tính theo công thức:

$$S_{th}^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i^0 - \overline{Y^0})^2$$

$$S_{th} = \sqrt{\frac{\sum_{i=1}^m (Y_i^0 - \overline{Y^0})^2}{m-1}}$$

trong đó m là số thí nghiệm ở tâm phương án.

- Sự có nghĩa của hệ số hồi quy được kiểm định theo tiêu chuẩn Student:

$$t_i = \frac{|b_i|}{S_{b_i}}$$

b_i : là hệ số thứ i trong phương trình hồi quy.

S_{b_i} : Độ lệch chuẩn của hệ số thứ i.

$$S_{b_i} = \frac{S_{th}}{\sqrt{N}}$$

➤ Phương sai tái hiện: $S_{th}^2 = 4.11$

Để kiểm định ý nghĩa các hệ số, tác giả tính các hệ số t_j (theo công thức trang 5), thu được kết quả sau:

t_0	t_1	t_2	t_3
21.3746	7.35263	5.42122	9.56644

Tra bảng phân phối phân vị Student với mức ý nghĩa $p = 0.05$, $f = N_0 - 1 = 2$ ta có $t_{0.05}(2) = 4.3$. Vậy các hệ số t_j đều lớn hơn $t_{0.05}(2)$ nên các hệ số của phương trình hồi qui đều có nghĩa.

Phương trình hồi qui có dạng sau:

$$\hat{Y}_L = 15.3275 - 5.2725X_1 + 3.8875X_2 + 6.86X_3$$

➤ Kiểm định sự tương thích của phương trình hồi qui với thực nghiệm:

STT	\hat{Y}_L	Y_i	$Y_i - \hat{Y}_L$	$(Y_i - \hat{Y}_L)^2$
1	20.81	12.35	-8.46	71.5716
2	7.09	8.87	1.78	3.1684
3	13.03	12.08	-0.95	0.9025
4	-0.69	6.92	7.61	57.9121

5	31.35	42.13	10.78	116.208
6	17.63	13.51	-4.12	16.9744
7	23.57	22.19	-1.38	1.9044
8	9.85	4.57	-5.28	27.8784

Phương sai dư (theo công thức trang 5): $S_{du}^2 = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N - L}$

(N là số thí nghiệm, L là hệ số ý nghĩa)

Ta có: $S_d^2 = 74.13$

Tiêu chuẩn Fisher:

$$F = S_d^2 / S_{th}^2 = 74.13 / 4.1 = 18.08$$

Tra bảng phân vị phân bố Fisher với $p = 0.05$; $f_1 = N-L = 4$; $f_2 = N_0-1 = 2$;
ta có:

$F_{1-p} = F_{0.05}(4,2) = 19.3$. Vậy $F < F_{0.05}(4;2)$. Phương trình hồi qui tương thích với thực nghiệm.

*Tối ưu hoá thực nghiệm bằng phương pháp leo dốc để tìm giá trị Y_{min} .

2. NHẬN XÉT

- Tác giả xác định hàm mục tiêu, bài toán tối ưu và phương án qui hoạch trực giao cấp 1 là phù hợp. Các số liệu được tác giả tính toán hầu như không sai lệch so với các số liệu đã được tính toán lại.

- Tuy nhiên, các số liệu thực nghiệm (Y_i) biến thiên bất hợp lý (không theo qui luật tuyến tính). Đồng thời, ba giá trị Y_0 của thí nghiệm tại tâm sai lệch nhau quá nhiều và khác rất xa so với hệ số B_0 . (Vì nếu các số liệu thực nghiệm đáng tin cậy và tính toán chính xác thì Y_0 phải xấp xỉ B_0) và giá trị trung bình của chúng là 7.503. Các công thức và phép tính được kiểm tra lại là đúng, vì vậy có thể nói các số liệu thực nghiệm Y_i chưa được chính xác.

- Có thể thực hiện bài toán tối ưu với hệ số tương tác.

5.2. Bài toán 2.

1. Nghiên cứu Ảnh hưởng của một số yếu tố công nghệ đến quá trình chiết tách anthocyanin

1.1. Ảnh hưởng của tỷ lệ dung môi

Nghiên cứu ảnh hưởng của tỷ lệ dung môi đến hàm lượng và độ màu anthocyanin thu được. Người N/C làm 6 thí nghiệm trong các điều kiện như sau:

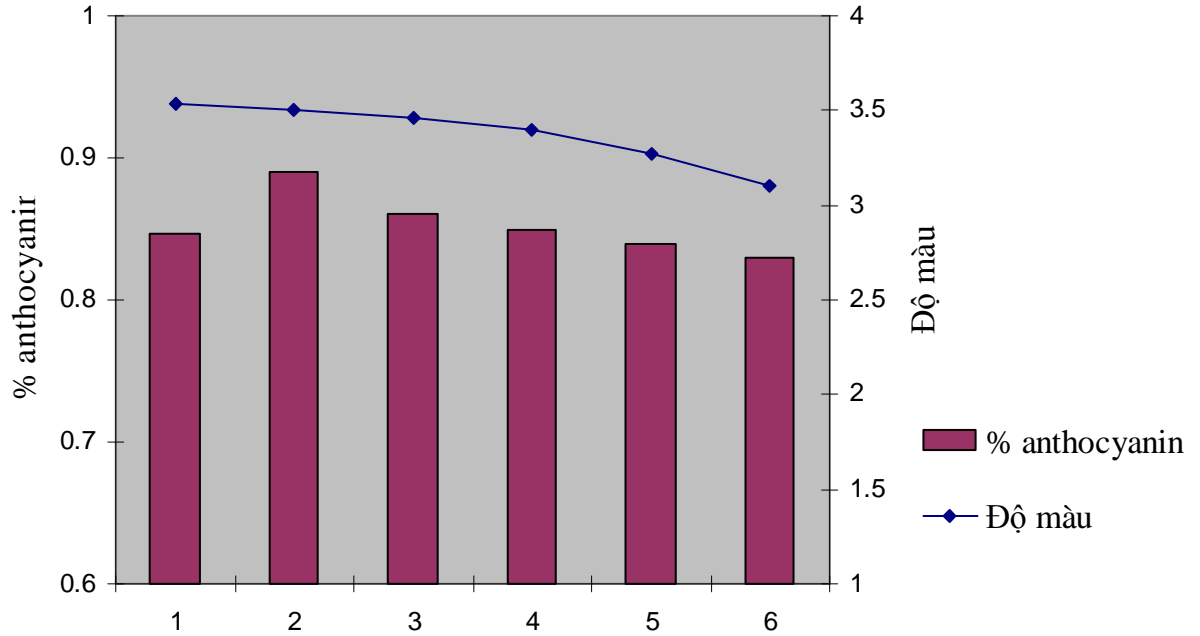
- Nhiệt độ chiết: 30°C

- Thời gian chiết: 45 phút

- Chiết trong hệ dung môi có tỷ lệ dung môi nước: ethanol thay đổi như bảng 1.

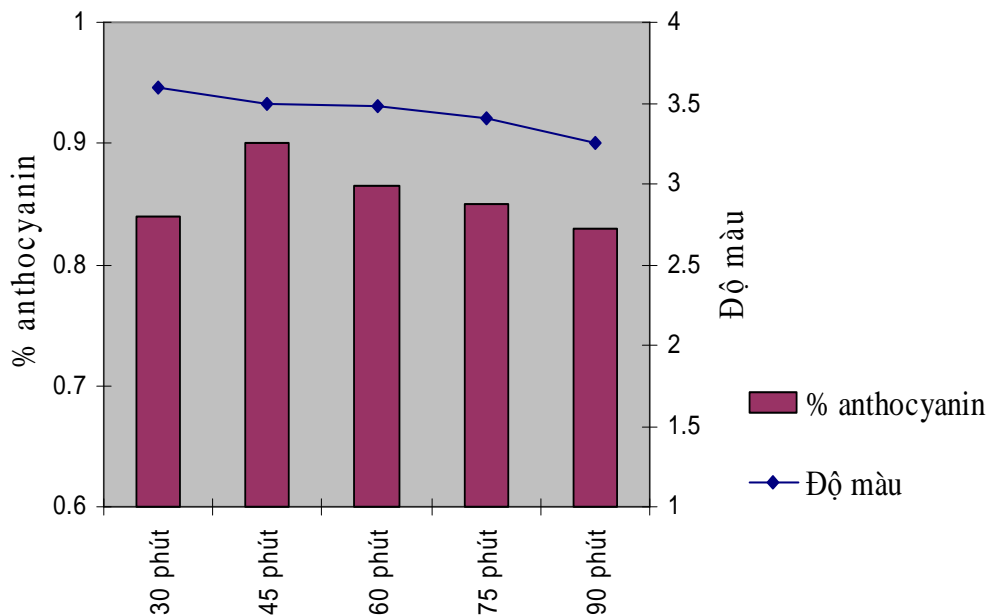
Bảng 1. Các thông số ban đầu và kết quả thí nghiệm

S TT	% V _{nước}	% V _{ethanol}	Tỷ lệ nước / ethanol	Hàm lượng % anthocyanin,	Độ màu
1	80	20	4/1	0,827	3,53
2	70	30	7/3	0.890	3,50
3	60	40	3/2	0.870	3,46
4	50	50	1/1	0,857	3,40
5	40	60	2/3	0,845	3,37
6	30	70	3/7	0,840	3.34



Hình 1: Biểu đồ biểu diễn ảnh hưởng của tỷ lệ dung môi thời gian chiết và độ màu anthocyanin

- Tiến hành 5 thí nghiệm trong các điều kiện sau:
- Nhiệt độ chiết: 30⁰C
 - Chiết trong dung môi có tỷ lệ nước: ethanol là 7/3
 - Thời gian thay đổi từ 30- 90 phút.

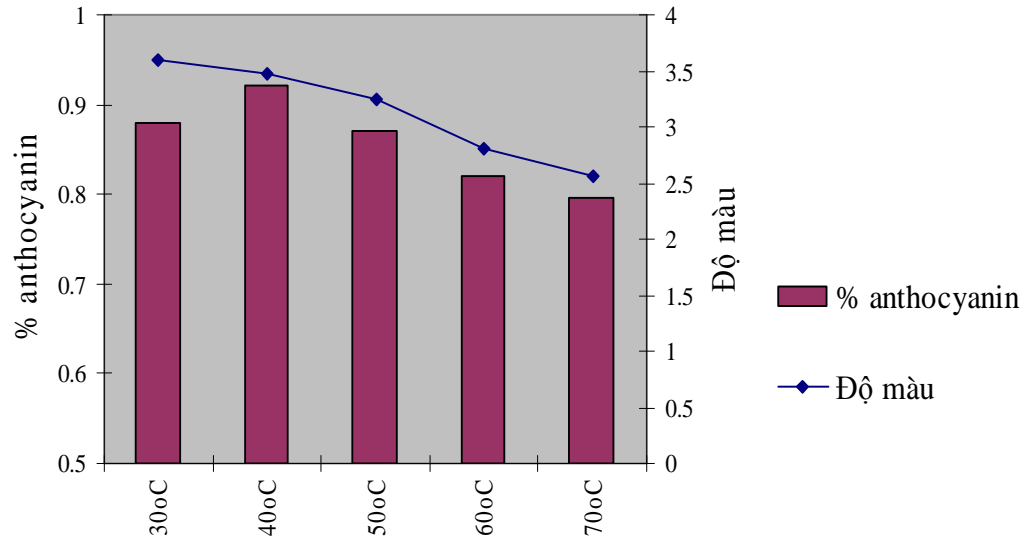


Hình 2: Biểu đồ biểu diễn ảnh hưởng của thời gian chiết đến hàm lượng và độ màu anthocyanin thu được

1.3. Ảnh hưởng của nhiệt độ

Nghiên cứu ảnh hưởng của nhiệt độ, tiến hành 5 thí nghiệm trong cùng điều kiện:

- Chiết trong hệ dung môi có tỷ lệ nước: ethanol là 7/3
- Thời gian chiết: 45 phút.
- Nhiệt độ chiết thay đổi từ 30°C ÷ 70°C . Các thông số cụ thể và kết quả được thể hiện trên biểu đồ hình (3).



Hình 3: Biểu đồ biểu diễn ảnh hưởng của nhiệt độ đến hàm lượng và độ màu anthocyanin

Qua nghiên cứu ảnh hưởng của một số yếu tố công nghệ đến khả năng thu nhận anthocyanin chúng tôi nhận thấy: Dung môi, tỷ lệ dung môi, thời gian chiết, nhiệt độ chiết đều ảnh hưởng đến khả năng chiết tách anthocyanin từ bắp cải tím. Ứng với mỗi điều kiện khác nhau chúng tôi thu được anthocyanin có hàm lượng và độ màu khác nhau. Từ các kết quả nghiên cứu chúng tôi chọn được miền khảo sát thích hợp của các yếu tố công nghệ cho các nghiên cứu tiếp theo như sau:

- Chiết trong hệ dung môi có tỷ lệ nước: ethanol dao động từ 7/3 ÷ 1/1
- Nhiệt độ chiết từ 30 ÷ 40°C
- Thời gian chiết trong khoảng 45 ÷ 75 phút

2. Tối ưu hoá điều kiện chiết tách anthocyanin có độ màu cao từ bắp cải tím

Với mục đích của đề tài là thu nhận và sử dụng chất màu anthocyanin, chúng tôi tiến hành tối ưu hoá điều kiện chiết tách trong khuôn khổ bài toán tối ưu đa mục tiêu để thu nhận chất màu anthocyanin có hàm lượng và độ màu cao nhất.

2.1 Chọn các yếu tố ảnh hưởng

Trong quá trình chiết tách anthocyanin phải chịu tác động của nhiều yếu tố công nghệ, song ở đây chúng tôi chọn 3 yếu tố đã được thăm dò ở phần trên:

- Z_1 : Nhiệt độ chiết, $^{\circ}\text{C}$
- Z_2 : Thời gian chiết, phút
- Z_3 : Tỷ lệ nước trong hệ dung môi, %
- Y_1 : Hàm lượng anthocyanin, %
- Y_2 : Độ màu

Phương trình biểu diễn mối quan hệ có dạng:

$$Y_1 = f(Z_1, Z_2, Z_3)$$

$$Y_2 = (Z_1, Z_2, Z_3)$$

$$Y_1 \longrightarrow \text{Max}$$

$$Y_2 \longrightarrow \text{Max}$$

Y_1 là hàm mục tiêu hàm lượng Y_2 Hàm mục tiêu độ màu.

2.2 Các bước thực hiện bài toán quy hoạch

2.2.1 Chọn phương án quy hoạch

Để xác định hướng đi của đề tài và nhanh chóng tiến tới miền tối ưu chúng tôi chọn phương án quy hoạch trực giao cấp I (TYT 2^k) thực nghiệm yếu tố toàn phần 2 mức, k yếu tố ảnh hưởng.

Phương trình hồi qui có dạng:

$$Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3 \quad (1)$$

Trong đó:

b_0 : Hệ số hồi qui.

b_1, b_2, b_3 : Hệ số tuyến tính

b_{12}, b_{23}, b_{13} : Hệ số tương tác đôi

b_{123} : Hệ số tương tác ba

Mỗi hệ số b đặc trưng cho ảnh hưởng của các yếu tố đến quá trình chiết tách.

2.2.2. Tổ chức thí nghiệm trực giao cấp I

Theo [1] số thí nghiệm trong phương án là $2^k = 8$, ($k = 3$) và điều kiện thí nghiệm được ghi ở bảng (2).

Bảng 2 : Điều kiện thí nghiệm được chọn:

Các mức	Các yếu tố ảnh hưởng		
	$Z_1, ^\circ\text{C}$	$Z_2, \text{phút}$	$Z_3, \%V$
Mức trên (+1)	40	75	70
Mức cơ sở (0)	35	60	60
Mức dưới (-1)	30	45	50
Khoảng biến thiên	5	15	10

Từ cách chọn phương án và điều kiện thí nghiệm, chúng tôi xây dựng ma trận thực nghiệm theo biến mã và tiến hành thí nghiệm theo ma trận.

Kết quả được ghi ở bảng (3)

Bảng 3. Ma trận thực nghiệm trực giao cấp I, $k=3$, và kết quả

STT	Biến mã							Y_1	Y_2
	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$		
1	+	+	+	+	+	+	+	0,975	3,433
2	-	+	+	-	-	+	-	1,102	4,525
3	+	-	+	-	+	-	-	0,849	4,255
4	-	-	+	+	-	-	+	1,109	4,987
5	+	+	-	+	-	-	-	0,854	2,007
6	-	+	-	-	+	-	+	0,717	3,988
7	+	-	-	-	-	+	+	0,944	3,205
8	-	-	-	+	+	+	-	0,813	3,767
T_1	0	0	0	0	0	0	0	0,915	3,834
T_2	0	0	0	0	0	0	0	0,935	3,991
T_3	0	0	0	0	0	0	0	0,955	3,773

Trong đó:

- x_1 : Nhiệt độ chiết, $^\circ\text{C}$
- x_2 : Thời gian chiết, phút
- x_3 : Tỷ lệ nước trong hệ dung môi, %
- Y_1 : Hàm mục tiêu hàm lượng anthocyanin, %
- Y_2 : Hàm mục tiêu độ màu.

2.2.3 Xây dựng mô tả toán học cho hàm mục tiêu hàm lượng anthocyanin

a) *Chọn phương trình hồi qui:*

Phương trình hồi qui được chọn theo phương trình (1) ở mục 2.2.1.

Các hệ số $b_1, b_2, b_3 \dots b_{123}$ được tính theo số liệu thực nghiệm hàm mục tiêu hàm lượng anthocyanin (Y_1).

b) *Tính hệ số b:*

Vì phương án được chọn là quy hoạch trực giao, theo [1] các hệ số b_j trong phương trình hồi qui (1) được xác định theo công thức sau:

$$\left. \begin{aligned} b_j &= \frac{1}{N} \sum_{u=1}^N (x_{ju} y_u) && \text{với : } j = (\bar{1}, k) \\ b_{ij} &= \frac{1}{N} \sum_{u=1}^N (x_{iu} x_{ju}) y_u && i \neq j = \bar{(1, k)} \\ b_{ijk} &= \frac{1}{N} \sum_{u=1}^N (x_{iu} x_{ju} x_{ku}) y_u && i \neq j \neq k \neq \bar{(1, k)} \end{aligned} \right\} \quad (2)$$

Từ số liệu thực nghiệm bảng (3), áp dụng các công thức (2) ta tính được các hệ số b:

$$\begin{aligned} b_0 &= 0,9208 && b_{12} = 0,017 \\ b_1 &= -0,07 && b_{13} = -0,0255 \\ b_2 &= 0,0488 && b_{23} = -0,019 \\ b_3 &= 0,088 && b_{123} = 0,0155 \end{aligned}$$

c) *Kiểm định mức ý nghĩa của các hệ số b trong phương trình 3.1*

Các hệ số được kiểm định theo tiêu chuẩn Student (t)

$$t_j = \frac{|b_j|}{S_{b_j}} \quad (3)$$

So sánh t_j với $t_p(f)$. Trong đó: - $t_p(f)$ là chuẩn student tra bảng ứng với xác suất tin cậy p và bậc tự do f, $f = n_0 - 1$.

b_j : là hệ số trong phương trình hồi quy đã chọn.

S_{b_j} là độ lệch của các hệ số b_j

Nếu $t_j > t_p(f)$ thì hệ số b_j có nghĩa.

Nếu $t_j < t_p(f)$ thì hệ số b_j bị loại khỏi phương trình.

Để kiểm định theo chuẩn Student (t) ta thay hệ số b_j, S_{b_j} vào công thức (3.3) ta có các giá trị t_j :

$$\begin{aligned} t_0 &= 130,21 && t_{12} = 2,510 \\ t_1 &= 10,01 && t_{13} = 3,606 \end{aligned}$$

$$t_2 = 6,894$$

$$t_{23} = 2,687$$

$$t_3 = 12,445$$

$$t_{123} = 2,192$$

Tra bảng tiêu chuẩn Student ta có $t_p(f_{th}) = t_{0,05}(2) = 4,3$

Do $t_{12} < t_p(f_{th})$, $t_{13} < t_p(f_{th})$, $t_{23} < t_p(f_{th})$, $t_{123} < t_p(f_{th})$ nên các hệ số b_{12} , b_{13} , b_{23} , b_{123} loại ra khỏi phương trình. Phương trình động học có dạng:

$$\hat{y}_1 = 0,9208 - 0,07x_1 + 0,04875x_2 + 0,088x_3 \quad (4)$$

d) *Kiểm định sự phù hợp của phương trình hồi qui với thực nghiệm*

Sự tương thích của phương trình với thực nghiệm được kiểm định

theo tiêu chuẩn Fisher (F).

$$F = \frac{S_{du}^2}{S_{th}^2} = 7,8406$$

Tra bảng tiêu chuẩn Fisher ta được $F_{1-p}(f_1, f_2) = F_{0,95}(4,2) = 19,3$

Ta có $F < F_{1-p}$. Vậy mô hình toán học đã chọn phù hợp với thực nghiệm.

2.2.4 Tối ưu hoá thực nghiệm để thu được hàm lượng anthocyanin cao nhất

a) *Tính các bước chuyển động δ_j :*

Từ mức cơ sở Z_j^0 , và phương trình hồi qui tuyến tính đối với hàm mục tiêu hàm lượng chúng tôi tính bước chuyển động δ_j ($j = 1, 2, 3$) cho mỗi yếu tố. Kết quả được ghi ở bảng 4.

Bảng 4. Kết quả tính bước chuyển động δ_j của các yếu tố

Các mức	Các yếu tố ảnh hưởng		
	$Z_1, ^\circ\text{C}$	$Z_2,$ phút	$Z_3,$ %
Mức cơ sở	35	60	60
Khoảng biến thiên (Δ_j)	5	15	10
Hệ số b_j	-0,070	0,048	0,088
$b_j\Delta_j$	-0,350	0,731	0,88
Bước chuyển động (δ_j)	-1,980	4,150	5
Làm tròn	-2	4	5

Theo bảng số liệu (4) ta có : $\Delta_3 b_{3_{\max}} = 0,88$, theo tài liệu [1]

Chọn bước chuyển động $\delta_3 = 0,5 \cdot \Delta_3 = 0,5 \cdot 10 = 5$.

Các bước chuyển động của yếu tố x_1 , x_2 được tính:

$$\delta_1 = \delta_3 \frac{b_1 \Delta_1}{b_3 \Delta_3} = -1,98$$

$$\delta_2 = \delta_3 \frac{b_2 \Delta_2}{b_3 \Delta_3} = 4,1$$

b) *Tổ chức thí nghiệm leo dốc:*

Từ kết quả các bước chuyển động δ_j ở bảng (4), chúng tôi tổ chức thí nghiệm leo dốc và điểm xuất phát từ tâm thực nghiệm.

Thí nghiệm theo hướng đã chọn, kết quả được biểu diễn ở bảng 5.

Bảng 5: Kết quả thí nghiệm theo hướng leo dốc

Yếu tố TN	$Z_1, ^\circ\text{C}$	$Z_2,$ phút	$Z_3,$ %	$y_1, \%$	y_2
1(Tn tại tâm)	35	60	60	0,927	
2	33	64	65	0.962	
3	31	68	70	0,985	
4	29	72	75	1,113	4,720
5	27	76	80	0,997	

Nhìn vào bảng 5, kết quả thí nghiệm tốt nhất thí nghiệm thứ tư. Tại nhiệt độ chiết 29°C , thời gian chiết 72 phút, tỷ lệ nước trong hệ dung môi là 75% chúng tôi thu được hàm lượng anthocyanin cao nhất là 1,113 %. Tại thí nghiệm này, chúng tôi xác định độ màu của anthocyanin là: 4,720. Đây chưa phải là độ màu thu được cao nhất. Vì thế, chúng tôi tiến hành tìm điều kiện chiết tách tối ưu để thu được anthocyanin có độ màu cao.

2.2.5 Xây dựng mô tả toán học với hàm mục tiêu độ màu.

a) *Chọn phương trình hồi qui:*

Phương trình hồi qui được chọn theo phương trình (1) ở mục 2.2.1. Các hệ số $b_1, b_2, b_3 \dots b_{123}$ được tính theo số liệu thực nghiệm hàm mục tiêu độ màu (Y_2).

b) *Kiểm tra mức ý nghĩa của hệ số b trong phương trình hồi qui:*

Sau khi kiểm tra mức ý nghĩa của các hệ số b ta có: $t_{13} < t_p(f_{th})$, $t_{23} < t_p(f_{th})$, $t_{123} < t_p(f_{th})$ nên các hệ số b_{13} , b_{23} , b_{123} bị loại ra khỏi phương trình.

Phương trình hồi qui có dạng:

$$\hat{y} = 3,7709 - 0,55x_1 - 0,2826x_2 + 0,5291x_3 - 0,2224x_1x_2 \quad (5)$$

c) *Kiểm định sự phù hợp của phương trình hồi qui với thực nghiệm:*

Các bước kiểm tra được trình bày ở phụ lục 6.

Sau khi kiểm tra ta có phương trình hồi qui (5) phù hợp với thực nghiệm.

2.2.6: Tối ưu hoá thực nghiệm để thu được anthocyanin có độ màu cao.

Sau khi kiểm tra phương trình hồi quy đã phù hợp với thực nghiệm, chúng tôi tiến hành tối ưu hoá thực nghiệm bằng phương pháp leo dốc để thu được anthocyanin có độ màu cao.

* Tính bước chuyển động của các yếu tố

Cũng từ mức cơ sở Z_j và phương trình hồi qui đối với hàm mục tiêu độ màu. Chúng tôi tính bước chuyển động δ_j ($j = 1, 2, 3$) tương tự như mục 2.2.3.

Kết quả được thể hiện ở bảng 6.

Bảng 6: Tính bước chuyển động δ_j của các yếu tố

Các mức	Các yếu tố		
	$Z_1, ^\circ\text{C}$	$Z_2,$ phút	$Z_3, \%$
Mức cơ sở	35	60	60
Khoảng biến thiên (Δ_j)	5	15	10
Hệ số b_j	-0,55	- 0,2826	0,5291
$b_j\Delta_j$	-2,57	- 4,239	5,29
Bước chuyển động (δ_j)	-2,07	-3,2	4
Làm tròn	-2	-3	4

*Tổ chức thí nghiệm leo dốc cho hàm mục tiêu độ màu:

Từ bước chuyển động δ_j ($j = 1, 2, 3$) ở bảng 3.10 chúng tôi tổ chức thí nghiệm theo hướng leo dốc, với điểm xuất phát là mức cơ sở Z_j^0 (Z_1^0, Z_2^0, Z_3^0). Chúng tôi thực

hiện thí nghiệm theo hướng đã chọn. Kết quả biểu diễn ở bảng 7.

Bảng 7: Kết quả thí nghiệm theo hướng leo dốc

Yếu tố TN	$Z_1, ^\circ\text{C}$	$Z_2,$ phút	$Z_3,$ %	y_2	y_1
1(Tn tại tâm)	35	60	60	3,991	
2	33	57	64	4,403	
3	31	54	68	4,927	
4	29	51	72	5,001	0,975
5	27	48	75	4,872	

Từ kết quả thực nghiệm theo hướng leo dốc ở bảng 7. Chúng tôi nhận thấy tại thí nghiệm thứ tư cho kết quả tốt nhất. Độ màu đạt giá trị 5,001, với điều kiện công nghệ: nhiệt độ 29°C , thời gian 51 phút, tỷ lệ nước trong hệ dung môi 72%. Tại thí nghiệm này, chúng tôi xác định hàm lượng anthocyanin là 0,975.

2.3. Tối ưu hoá hàm đa mục tiêu bằng phương pháp chập tuyến tính.

Quá trình chiết tách chất màu anthocyanin có độ màu cao từ bắp cải tím được đặc trưng bởi hai phương trình (4), (5). Hai phương trình này thể hiện sự tác động của các yếu tố công nghệ đến hàm lượng và độ màu anthocyanin thu được.

Khi có sự thay đổi của bộ số liệu (x_1, x_2, x_3) trong bảng ma trận thực nghiệm (56) thì cho các giá trị thực nghiệm y_1, y_2 khác nhau và $y_{1\max}, y_{2\max}$ cũng khác nhau.

Với mục đích thu nhận chất màu anthocyanin có hàm lượng và độ màu cao, nhiệm vụ của chúng tôi phải tối ưu hoá hàm đa mục tiêu để tìm giải pháp

công nghệ thực tiễn tốt cho cả hai hàm mục tiêu, đồng thời nâng cao tính toàn diện và tính thuyết phục cho kết quả thu được.

Thực tế không thể có một nghiệm chung cho cả hai quá trình để đạt được y_{1max} , y_{2max} mà chỉ tìm được nghiệm thỏa hiệp (x_1, x_2, x_3) để các giá trị y_1, y_2 nằm gần y_{1max}, y_{2max} . Để tìm được nghiệm thỏa hiệp chúng tôi sử dụng phương pháp chập tuyến tính :

$$y_L = \alpha_1 y_1 + \alpha_2 y_2$$

Trong đó :

- α_1 là hệ số quan trọng ứng với hàm mục tiêu hàm lượng (y_1)
- α_2 là hệ số quan trọng ứng với hàm mục tiêu độ màu (y_2)

Với mục đích thu nhận chất màu anthocyanin để ứng dụng làm chất chỉ thị trong hoá phân tích. chúng tôi ưu tiên cho hàm mục tiêu hàm lượng.

Chọn: $-\alpha_1 = 0,6, \alpha_2 = 0,4$

Ta có phương trình hàm đa mục tiêu : $y_L = 0,6y_1 + 0,4y_2$

Các hệ số của phương trình hồi quy được tính theo bảng 8

Bảng 8. Tính hệ số của phương trình hồi quy

Hệ số b	y_1	y_2	y_L
b_0	0,928	3,7709	2,06
b_1	-0,07	-0,55	-0,262
b_2	0,04875	-0,2826	-0,0837
b_3	0,088	0,5291	0,2644
b_{12}		-0,2223	-0,22

Ta có phương trình hồi quy:

$$y_L = 2,06 - 0,262x_1 - 0,08379x_2 + 0,2644x_3 - 0,222x_1x_2 \quad (6)$$

Tiến hành tối ưu hóa hàm đa mục tiêu để tìm giải pháp công nghệ thực tiễn phù hợp.

*Tính các bước chuyển động δ_j : cho HMT y_L

Cũng tương tự như mục 2.2.3, chúng tôi tính bước chuyển động δ_j ($\delta_1, \delta_2, \delta_3$) cho các yếu tố ảnh hưởng. Kết quả được biểu diễn ở bảng 9.

Bảng 9. Tính bước chuyển động của các mức yếu tố

Các mức	Các yếu tố		
	$Z_1, ^\circ\text{C}$	$Z_2,$ phút	$Z_3, \%$
Mức cơ sở	35	60	60
Khoảng biến thiên (Δ_j)	5	15	10
Hệ số b_j	- 0,262	- 0,08397	0,2644
$b_j \Delta_j$	-1,31	- 1,427	3
Bước chuyển động (δ_j)	-1,48	- 1,427	3
Làm tròn	-1,5	-1,5	3

* Tổ chức thí nghiệm leo dốc cho hàm mục tiêu Y_L :

Bảng 10: Kết quả thí nghiệm theo hướng leo dốc của hàm chấp Y_L

K TN	Các yếu tố ảnh hưởng			$Y_1, \%$	Y_2	Y_L
	$Z_1, ^\circ\text{C}$	$Z_2,$ phút	$Z_3, \%$			
1 (TN tại tâm)	35	60	60	0,927	3,991	2,149
2	33,5	58,5	63	0,943	4,403	2,287
3	32	57	66	0,972	4,911	2,548
4	30,5	55,5	69	0,983	4,952	2,553
5	29	54	72	1,110	4,967	2,656
6	28,5	52,5	75	0,915	4,500	2,397

Nhìn vào bảng 10, tại thí nghiệm thứ 5 hàm chấp y_L đạt giá trị lớn nhất $y_{L\max} = 2,656$. So sánh với các thí nghiệm leo dốc với từng hàm mục tiêu ta được

PA.TN	$Z_1, ^\circ\text{C}$	$Z_2,$	$Z_3, \%$	$y_1, \%$	y_2	y_L
-------	-----------------------	--------	-----------	-----------	-------	-------

		phút				
Theo % Anth	29	75	75	1,113	4,72	2,556
Theo độ màu	29	51	72	0,975	5,001	2,585
Theo hàm chạp	29	54	72	1,110	4,967	2,656

Từ bảng trên chúng tôi đã tìm được điều kiện tốt để chiết chất màu anthocyanin từ búp cải tím trong môi trường trung tính là:

Nhiệt độ chiết 29 °C

Thời gian chiết là 54 phút

Hệ dung môi nước -ethanol là 72-28

5.3. Bài toán 3.

KẾT QUẢ NGHIÊN CỨU ẢNH HƯỞNG CỦA NHIỆT ĐỘ VÀ NỒNG ĐỘ KIỀM ĐẾN QUÁ TRÌNH TÁCH TẠP CHẤT RA KHỎI XƠ SỢI XENLULO

(Trích một phần đề tài luận văn Thạc sĩ của CN Nguyễn Bá Trung – khoa Hóa, trường Đại học Sư phạm ĐHQĐHN)

I. KẾT QUẢ NGHIÊN CỨU TỪ ĐỀ TÀI

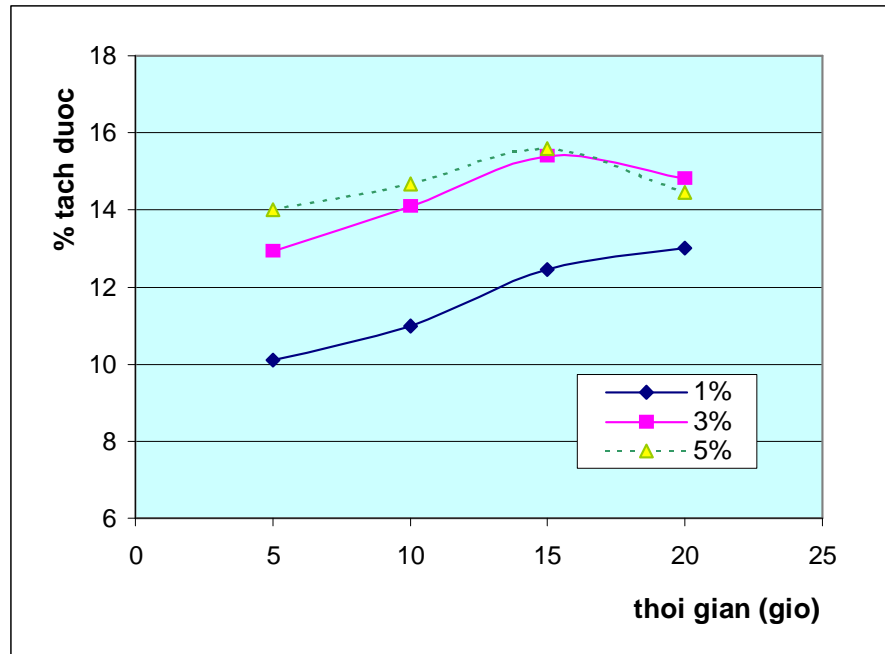
1. Nghiên cứu ảnh hưởng của nhiệt độ và nồng độ kiềm đến quá trình tách tạp chất ra khỏi xơ sợi xenlulo.

1.1. Ảnh hưởng của thời gian và nồng độ NaOH đến lượng tạp chất tách ra.

Gai bẹ sau khi đã được tách sơ bộ, dùng tay tước gai bẹ khô ra thành những sợi mỏng, cột lại thành bó, các bó có khối lượng xấp xỉ nhau khoảng 15 gam. Ngâm các bó sợi vào dung dịch NaOH ở các nồng độ nghiên cứu trong thời gian tương ứng. Kết quả thu được ở bảng sau:

Bảng 1.1: Kết quả khảo sát ảnh hưởng của thời gian và nồng độ đến lượng tạp chất tách ra

Thời gian C%	5 giờ	10 giờ	15 giờ	20 giờ
1% NaOH	10.098	10.978	12.46	13.01
3% NaOH	12.92	14.1	15.4	14.81
5% NaOH	14.01	14.67	15.598	14.46



Hình 1.1. Ảnh hưởng thời gian và nồng độ NaOH đến lượng tạp chất tách ra.

➤ **Nhận xét:** Từ 3 đường biểu diễn ở nồng độ 1%, 3% và 5% ta nhận thấy rằng ở nồng độ kiềm 1% thì hiệu quả tách là không cao. Khi tăng nồng độ kiềm lên 3% và 5% thì hiệu quả tách tăng lên nhiều. Tuy nhiên hiệu quả tách trong khoảng nồng độ từ 3 ÷ 5% là không khác nhau lắm trong khoảng thời gian từ 10 ÷ 15h.

Điều này có thể được giải thích như sau: Ở nồng độ quá loãng 1%, thời gian ngắn ban đầu chưa đủ để hoà tan các tạp chất bao bọc bên ngoài nên hiệu quả của quá trình tách không cao. Sau thời gian từ 10 –15 giờ, các lớp bên ngoài đã bị hoà tan nên tạo điều kiện thuận lợi cho NaOH thâm nhập vào bên trong để hoà tan hemixenlulo, lignin và các chất có phân tử lượng thấp khác có trong các bó sợi.

Như vậy qua đồ thị cho ta thấy rằng cả thời gian và nồng độ đều có ảnh hưởng rõ nét đến quá trình tách tạp chất ra khỏi xơ sợi gai.

1.1.2. Tối ưu hoá thực nghiệm quá trình tách tạp chất ra khỏi sợi gai.

Để tiến tới miền tối ưu, chúng tôi chọn phương án thực nghiệm yếu tố toàn phần. Hai yếu tố ảnh hưởng đến quá trình là nồng độ (Z_1) và thời gian ngâm

(Z_2). Hàm mục tiêu cần đạt được là lượng tạp chất tách ra khỏi sợi là lớn nhất hay nói cách khác hiệu quả tách là cao nhất.

Để quy hoạch thực nghiệm toàn phần, chúng tôi đã tiến hành bố trí thí nghiệm thay đổi đồng thời các yếu tố, mỗi yếu tố được tiến hành ở 3 mức: mức trên, mức dưới và mức cơ sở để thí nghiệm ở tâm phương án

Mức trên, mức dưới, khoảng biến thiên được trình bày ở bảng 1.2, ma trận quy hoạch thực nghiệm được trình bày ở bảng 1.3.

Bảng 1.2. Các mức thí nghiệm.

	Mức dưới	Mức cơ sở	Mức trên	Khoảng biến thiên (Δ)
Z_1 (C% NaOH)	3	4	5	1
Z_2 (thời gian ngâm)	10	12.5	15	2.5

➤ Lập ma trận quy hoạch:

Với 2 yếu tố nhiệt độ và nồng độ ($k = 2$), mỗi yếu tố có hai mức là mức trên và mức dưới. Vậy số thí nghiệm được tiến hành là $N = 2^2 = 4$ thí nghiệm.

Phương án tiến hành trình bày ở bảng sau:

Bảng 1.3. Ma trận quy hoạch thực nghiệm.

N	Các yếu tố theo tỉ lệ thực		Các yếu tố theo tỉ lệ mã hoá				Giá trị đo được
	Z_1 (C%)	Z_2 (t)	X_0	X_1	X_2	X_1X_2	Y
1	5	15	1	1	1	1	15.598
2	3	10	1	-1	-1	1	14.1
3	5	10	1	1	-1	-1	14.67
4	3	15	1	-1	1	-1	15.4
5	4	12.5	1	0	0	0	14.82
6	4	12.5	1	0	0	0	14.8
7	4	12.5	1	0	0	0	14.75

➤ Thiết lập phương trình hồi quy:

Tính hệ số hồi quy: Các hệ số hồi quy được tính theo công thức toán học (3), (4), (5). Từ số liệu thực nghiệm ta xác định được các giá trị b_0 , b_1 , b_2 như sau:

$$b_0 = 14,982 \quad ; \quad b_1 = 0,192 \quad ; \quad b_2 = 0,557 \quad ; \quad b_{12} = - 0,093$$

Với kết quả trên ta có phương trình hồi quy theo toán học:

$$Y = 14,982 + 0,192 X_1 + 0,557 X_2 - 0,093 X_1X_2$$

Để kiểm định ý nghĩa của hệ số hồi quy và sự tương thích của phương trình hồi quy với thực nghiệm, ta phải tìm phương sai tái hiện S_{th}^2 . Do vậy chúng ta phải làm thêm 3 thí nghiệm ở tâm phương án và thu được giá trị \bar{Y}^0 .

Từ công thức tính phương sai tái hiện, ta được: $S_{th}^2 = 0,0013$; $S_{th} = 0.03605551$

➤ **Kiểm định các hệ số có nghĩa của phương trình hồi quy:**

Sự có nghĩa của hệ số hồi quy được theo tiêu chuẩn Student: $t_i = \frac{|b_i|}{S_{b_i}}$

Bằng cách tính như trên ta thu được các giá trị t_i như sau:

$$t_0 = 828,833 \quad t_2 = 30,896$$

$$t_1 = 10,650 \quad t_{12} = 5,158$$

Tra bảng tiêu chuẩn Student ta có $t_{0,05}(2) = 4,303$.

Qua bảng số liệu trên ta thấy các hệ số b_0 , b_1 , b_2 , b_{12} là có ý nghĩa với độ tin cậy $P < 0,05$.

Vậy phương trình hồi quy mô tả ảnh hưởng của nồng độ kiềm và nhiệt độ lên quá trình tách tạp chất ra khỏi sợi gai bẹ có dạng sau:

$$Y_1 = 14,982 + 0,192 X_1 + 0,557 X_2 - 0,093 X_1X_2$$

Nhận xét: Sự có mặt của các giá trị b_1 , b_2 và b_{12} trong phương trình hồi quy cho thấy cả 2 yếu tố nồng độ và thời gian đều ảnh hưởng đến quá trình tách các chất có phân tử lượng thấp ra khỏi gai bẹ. Đặc biệt, sự có mặt của giá trị b_{12} chứng tỏ có sự tương tác qua lại giữa hai yếu tố nồng độ kiềm và thời gian ngâm.

Giá trị b_{12} bé hơn nhiều so với b_1 , b_2 và mang dấu âm, trong khi đó b_1 và b_2 đều mang dấu dương nên về mặt toán học cho ta có những nhận xét sau:

- Sự tương tác giữa hai yếu tố trong khoảng khảo sát là nhỏ, không đáng kể và có thể bỏ qua sự tương tác đó. Nếu vượt ra khỏi khoảng khảo sát, thì sự tương tác qua lại giữa hai yếu tố có thể ảnh hưởng đến lượng tạp chất được tách ra.

- Trong khoảng khảo sát muốn tăng hiệu quả tách tạp chất thì phải tăng nồng độ kiềm và thời gian ngâm. Tuy nhiên trong 2 yếu tố trên thì thời gian ngâm đóng vai trò quan trọng hơn do giá trị hệ số b_2 là lớn hơn so với b_1 .

➤ **Tối ưu hoá thực nghiệm.**

Chúng tôi tiến hành tối ưu hoá quá trình tách tạp chất ra khỏi sơ sợi gai bằng phương pháp thực nghiệm leo dốc nhất. Với các hệ số của phương trình hồi quy như trên, tăng nồng độ kiềm và thời gian thì hiệu quả tách tạp chất sẽ tăng, đến một giới hạn nào đó thì giá trị phần trăm tạp chất tách ra là cực đại và quá giới hạn đó thì giảm xuống.

Tối ưu hoá được thực hiện như sau:

Chọn bước nhảy của yếu tố x_1 là $\delta_1 = 0,05$. Dựa vào δ_1 ta tính được giá trị

δ_2 theo công thức sau :

$$\delta_1 = \delta_3 \frac{b_1 \Delta_1}{b_3 \Delta_3}$$

$$\delta_2 = \delta_3 \frac{b_2 \Delta_2}{b_3 \Delta_3}$$

được: $\delta_2 = 0,05 \cdot \frac{0,557 \cdot 2,5}{0,192 \cdot 1} = 0,36$ Từ các thông số đã tính được, chúng tôi tiến

hành thực nghiệm tối ưu hoá các mức cơ sở và bước nhảy như sau:

Bảng 1.4. Kết quả tối ưu phương trình hồi quy Y_1 mô tả ảnh hưởng của nồng độ và thời gian lên hàm lượng các chất tách ra.

<i>Các yếu tố</i>	<i>X_1 (nồng độ kiềm C%)</i>	<i>X_2 (thời gian ngâm)</i>
<i>Mức cơ sở</i>	4 %	12,5 giờ = 750 phút
<i>Bước nhảy δ</i>	$\delta_1 = 0,05$	$\delta_2 = 0,36$ giờ \approx 20 phút
Nồng độ NaOH	Thời gian ngâm	Y (lượng chất tách ra)

4,05	770 phút	14,3
4,10	790 phút	14,81
4,15 (TNTT)	810 phút	-
4,20 (TNTT)	830 phút	-
4,25	850 phút	15,34
4,30	870 phút	15,62
4,35	890 phút	15,2

Như vậy để hiệu quả tách tạp chất ra khỏi xơ sợi gai thì phải tiến hành ở các thông số sau ở nhiệt độ phòng.

<u>Các thông số tối ưu</u>	<u>Giá trị</u>
Nồng độ kiềm C%	4,3%
<i>Thời gian ngâm</i>	870 phút (14 giờ 30 phút)

II. NHẬN XÉT VÀ ĐÁNH GIÁ VỀ BÀI TOÁN TỐI ƯU HÓA TRONG ĐỀ TÀI

1. Về phương pháp nghiên cứu

3.1.1. Phương pháp quy hoạch thực nghiệm.

Trong đề tài trên, CN NBT đã sử dụng phương pháp quy hoạch thực nghiệm toàn phần. Nội dung của phương pháp như sau:

- Khảo sát ảnh hưởng của các yếu tố đến quá trình tách tạp chất ra khỏi sợi xenlulo, chọn ra các yếu tố ảnh hưởng nhất.
- Xây dựng ma trận quy hoạch thực nghiệm.
- Tổ chức thí nghiệm và lấy mẫu thống kê.
- Từ số liệu thu được, xây dựng mô hình thống kê thực nghiệm để mô tả quá trình.

- Kiểm định thống kê.
- Phân tích sự tương tác giữa các yếu tố ảnh hưởng đến quá trình xử lý sợi trong phạm vi đã chọn.

Để quy hoạch thực nghiệm toàn phần, tác giả đã tiến hành bố trí thí nghiệm thay đổi đồng thời các yếu tố. Mỗi yếu tố được tiến hành ở 3 mức: Mức trên; mức dưới; mức cơ sở để thí nghiệm ở tâm phương án. Xây dựng hàm mục tiêu cần đạt được (max hay min).

➤ Lập ma trận quy hoạch:

Với 2 yếu tố là nhiệt độ và nồng độ ($k = 2$), mỗi yếu tố có hai mức là mức trên và mức dưới. Vậy số thí nghiệm được tiến hành là $N = 2^2 = 4$ thí nghiệm.

Để tiện cho việc tính toán, ta chuyển từ hệ trực tự nhiên Z_1, Z_2 có thứ nguyên sang hệ trục không thứ nguyên mã hoá. Việc mã hoá được thực hiện dễ dàng nhờ chọn tâm của miền được nghiên cứu làm gốc toạ độ.

Trong hệ mã hoá không thứ nguyên ta có được:

Mức trên: - kí hiệu +1

Mức cơ sở: - kí hiệu 0

Mức dưới: - kí hiệu -1

Công thức chuyển từ hệ đơn vị thực qua đơn vị mã hoá không thứ nguyên:

$$X_j = \frac{Z_j - Z_j^0}{\Delta Z_j}; j = 1, \dots, k \quad (1)$$

$$\Delta Z_j = \frac{Z_j^{\max} - Z_j^{\min}}{2}; j = 1, \dots, k \quad (2)$$

➤ **Thiết lập phương trình hồi quy mô tả ảnh hưởng của các yếu tố đến quá trình nghiên cứu:**

Tính hệ số hồi quy: Các hệ số hồi quy được tính theo công thức toán học như sau:

$$b_0 = \frac{\sum_{i=1}^n Y_i}{N} \quad (3)$$

$$b_i = \frac{\sum_{i=1}^n X_{ij} Y_i}{N} \quad (4)$$

$$b_{ji} = \frac{\sum_{i=1}^n (X_j \cdot X_1)_i \cdot Y_i}{N} \quad (5)$$

Từ số liệu thực nghiệm trên, áp dụng các công thức (3), (4) và (5) xác định được giá trị b_0 , b_1 , b_2 và b_{12} .

Từ kết quả trên có phương trình theo toán học: $Y = b_0 + b_1X_1 + b_2X_2 + b_{12}X_1X_2$

➤ **Kiểm định ý nghĩa của hệ số hồi quy và sự tương thích của phương trình hồi quy với thực nghiệm.** Tìm phương sai tái hiện S_{th}^2 . Do vậy, phải làm thêm 3 thí nghiệm ở tâm phương án và thu được ba giá trị Y_i^0 và giá trị \bar{Y}^0 tại tâm.

Phương sai tái hiện được tính theo công thức:

$$S_{th}^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i^0 - \bar{Y}^0)^2 \quad (6)$$

$$S_{th} = \sqrt{\frac{\sum_{i=1}^m (Y_i^0 - \bar{Y}^0)^2}{m-1}} \quad (7)$$

trong đó m là số thí nghiệm ở tâm phương án.

- Sự có nghĩa của hệ số hồi quy được kiểm định theo tiêu chuẩn Student:

$$t_i = \frac{|b_i|}{S_{b_i}}$$

b_i : là hệ số thứ i trong phương trình hồi quy.

S_{b_i} : Độ lệch quân phương của hệ số thứ i .

$$S_{b_i} = \frac{S_{th}}{\sqrt{N}} \quad (8)$$

Tra bảng tiêu chuẩn Student ta có $t_{0,05}(2) = 4,303$.

Nếu giá trị $t_j > t_{0,05}(2) = 4,303$ thì các hệ số b_j là có nghĩa, ngược lại là không có nghĩa.

Xây dựng phương trình hồi quy với các hệ số b_j có nghĩa.

- *Kiểm định sự tương thích của phương trình hồi quy với thực nghiệm theo tiêu chuẩn Fisher.*

$$\text{Ta có } F = \frac{S_{du}^2}{S_{th}^2} \quad (9)$$

$$\text{với } S_{\text{du}}^2 = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N - L} \quad (10)$$

$$(N \text{ là số thí nghiệm, } L \text{ là hệ số ý nghĩa}) \quad Y_i = \sum_{i=1}^4 b_i X_i \quad (11)$$

Thay số vào tính toán ta được các giá trị $Y_1; Y_2; Y_3; Y_4$. So sánh với $F_{(1-p)(f_1, f_2)}$ với $P = 0,05$; f_1 là bậc tự do của phương sai tương thích ($f_1 = n - 1$); f_2 là bậc tự do của phương sai tái hiện ($f_2 = m - 1$). Tra bảng ta có được giá trị $F_{(0,95)(f_1, f_2)}$

Nếu $F < F_{\text{tb}}$ thì phương trình hồi quy tìm được là tương thích với thực nghiệm với mức ý nghĩa 95% , ngược lại là không tương thích.

1.3. Tối ưu hoá thực nghiệm.

Tiến hành tối ưu hóa bằng phương pháp thực nghiệm leo dốc nhất.

Vectơ $\hat{\text{grad}} y$ là một vectơ có chiều biểu thị sự biến thiên nhanh nhất của $\hat{\text{grad}} y(x)$, giá trị của $\hat{\text{grad}} y(x)$ thay đổi từ điểm này sang điểm khác trong không gian yếu tố.

Với mô hình tuyến tính bội k,

$$\hat{\text{grad}} y = \frac{\partial \hat{y}}{\partial x_1} \vec{i} + \frac{\partial \hat{y}}{\partial x_2} \vec{j} + \dots + \frac{\partial \hat{y}}{\partial x_k} \vec{k} \quad (12)$$

$$\text{hoặc } \hat{\text{grad}} y = b_1 \vec{i} + b_2 \vec{j} + \dots + b_k \vec{k} \quad (13)$$

Chuyển động theo $\hat{\text{grad}} y$ là chuyển động theo đường ngắn nhất đến điểm tối ưu, bởi vì hướng $\hat{\text{grad}}$ là hướng có độ nghiêng dốc nhất dẫn từ điểm đã cho đến điểm cực đại.

Trong trường hợp k yếu tố việc tính đường dốc nhất trên mặt đáp ứng được thực hiện như sau:

Chọn bước nhảy của yếu tố x_1 là δ_1 ; dựa vào δ_1 ta tính được δ_2 theo công

$$\text{thức:} \quad \delta_j = \delta_1 \cdot \frac{b_j \Delta_j}{b_i \Delta_i}$$

(14)

trong đó: δ_i là bước nhảy của yếu tố thứ i

b_i, b_j là hệ số hồi quy của các yếu tố tương quan

Δ_i, Δ_j là khoảng biến thiên của từng yếu tố tương ứng
thay số vào ta tính được các giá trị δ_j khác.

Từ các thông số đã tính được, tiến hành thực nghiệm tối ưu hoá các mức cơ sở với các bước nhảy đã tính sẵn trước. Chuyển động grad phải bắt đầu từ điểm 0 (mức cơ sở của mỗi yếu tố) và dừng lại khi tìm được điểm tối ưu nếu những hạn chế đặt vào các yếu tố làm cho chuyển động tiếp tục theo hướng grad không hợp lý nữa.

5.4. Bài toán 4.

Tối ưu hoá thực nghiệm quá trình chiết các hợp chất từ gỗ vang

Tiến hành tối ưu hóa thực nghiệm quá trình chiết bằng phương pháp chung nình với dung môi là H₂O. Để tiến tới miền tối ưu, chúng tôi chọn phương án thực nghiệm yếu tố toàn phần. Ba yếu tố ảnh hưởng đến quá trình là khối lượng gỗ (Z₁ -g), thể tích dung môi (Z₂ - ml) và thời gian chung (Z₃ - h). Hàm mục tiêu cần đạt được là độ hấp thụ quang lớn nhất. Mức trên, mức dưới, khoảng biến thiên được trình bày ở bảng 1.

Bảng 1. Các mức thí nghiệm.

Các yếu tố	Mức dưới	Mức cơ sở	Mức trên	Khoảng biến thiên (Δ)
Z ₁ (m – g)	5	7,5	10	2,5
Z ₁ (V-ml)	50	100	150	50
Z ₂ (thời gian - h)	4	7	10	3

Trong hệ mã hoá không thứ nguyên ta có được:

Mức trên: - kí hiệu +1

Mức cơ sở: - kí hiệu 0

Mức dưới: - kí hiệu -1

Lập ma trận quy hoạch: Với 3 yếu tố khối lượng gỗ, thể tích dung môi và thời gian chung, mỗi yếu tố có hai mức là mức trên và mức dưới. Vậy số thí nghiệm được tiến hành là $N = 2^3 = 8$ thí nghiệm. Phương án tiến hành, ma trận quy hoạch thực nghiệm được trình bày ở bảng 2.

Bảng 2. Ma trận quy hoạch thực nghiệm

TN	Các yếu tố theo tỉ lệ thực			Các yếu tố theo tỉ lệ mã hoá				Mật độ quang
	Z ₁	Z ₂	Z ₃	X ₀	X ₁	X ₂	X ₃	Y _i
1	10	150	10	1	1	1	1	0,1083
2	5	50	10	1	-1	-1	1	0,1009
3	10	50	10	1	1	-1	1	0,1085
4	5	150	10	1	-1	1	1	0,0855
5	10	150	4	1	1	1	-1	0,1078

6	5	50	4	1	-1	-1	-1	0,1092
7	10	50	4	1	1	-1	-1	0,1105
8	5	150	4	1	-1	1	-1	0,1047
9	7.5	100	7					0,1060
10	7.5	100	7					0,1072
11	7.5	100	7					0,1085

➤ **Thiết lập phương trình hồi quy:**

Phương trình hồi quy theo toán học (3):

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 \quad (3)$$

Các hệ số hồi quy b_j được tính theo công thức toán học như sau:

$$b_0 = \frac{\sum_{i=1}^n Y_i}{N} \quad (4)$$

$$b_i = \frac{\sum_{i=1}^n X_{ij} Y_i}{N} \quad (5)$$

$$b_{jl} = \frac{\sum_{i=1}^n (X_j \cdot X_l)_i \cdot Y_i}{N} \quad (6)$$

Tính hệ số hồi quy: Các hệ số hồi quy được tính theo công thức toán học (4), (5), (6). Từ số liệu thực nghiệm ta xác định được các giá trị b_0, b_i, b_{ij} như sau:

b_0	b_1	b_2	b_3	b_{12}	b_{23}	b_{31}
0.147	0.0046	-0.0031	-0.0039	0.019	-0.0008	0.0030

Với kết quả trên ta có phương trình hồi quy theo toán học:

$$Y = 0.1047 + 0.0046X_1 - 0.0031X_2 - 0.0039X_3 + 0.0019X_1X_2 - 0.0008X_2X_3 + 0.0030X_3X_1$$

Để kiểm định ý nghĩa của hệ số hồi quy và sự tương thích của phương trình hồi quy với thực nghiệm phải làm thêm 3 thí nghiệm ở tâm phương án (thí nghiệm 9, 10, 11 trong bảng 3.8) để xác định phương sai tái hiện theo công thức (2.9) ta có: $S_{th}^2 = 1,56 \cdot 10^{-6}$. Từ đó: $S_{th} = 0.0013$.

➤ **Kiểm tra sự có nghĩa của hệ số hồi quy được kiểm định theo tiêu chuẩn Student:**

$$t_i = \frac{|b_i|}{S_{b_i}} \quad (7)$$

Với S_{b_j} là độ lệch quân phương tính bằng: $S_{b_j} = \frac{S_{th}}{\sqrt{N}}$ (8)

S_{th} là phương sai tái hiện đối với 3 thí nghiệm tại tâm tính theo công thức:

$$S_{th}^2 = \frac{\sum_{u=1}^3 (y_u^0 - \bar{y}^0)^2}{3-1} \quad (9)$$

Theo công thức (7): $t_i = \frac{|b_i|}{S_{b_i}}$ thu được các giá trị t_i như sau:

t_0	t_1	t_2	t_3	t_{12}	t_{23}	t_{31}
166.828	7.332	4.941	6.197	2.969	1.275	3.782

Tra bảng tiêu chuẩn Student ta có $t_{0,05}(2) = 4,303$.

Qua các giá trị t_i cho thấy: $t_{12}, t_{23}, t_{31} < t_{0,05}$ vậy các hệ số $b_{12}; b_{23}$ không có ý nghĩa, các hệ số còn lại $b_0, b_1, b_2, b_3, b_{31}$ là có ý nghĩa với độ tin cậy $P=0,05$.

Từ đó phương trình hồi quy mô tả ảnh hưởng của các yếu tố lên quá trình chiết các hợp chất hóa học từ cây gõ vang có dạng sau:

$$Y = 0.147 + 0.0046X_1 - 0.0031X_2 - 0.0039X_3$$

➤ **Kiểm định sự tương thích của phương trình hồi quy với thực nghiệm theo tiêu chuẩn Fisher.**

➤ **Kiểm định sự tương thích của phương trình hồi quy với thực nghiệm theo tiêu chuẩn Fisher.**

$$\text{Ta có } F = \frac{S_{du}^2}{S_{th}^2} \quad (10)$$

$$\text{với } S_{du}^2 = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N-L} \quad (11)$$

(N là số thí nghiệm, L là hệ số ý nghĩa).

Các giá trị \hat{Y}_i tính được theo phương trình hồi qui: $\hat{Y}_i = \sum_{i=1}^4 b_i X_i$

$F_{tb} = F_{(1-p)(f_1, f_2)}$ với mức ý nghĩa 95% thì $p = 0,05$; f_1 là bậc tự do của phương sai tương thích ($f_1 = N - L$); f_2 là bậc tự do của phương sai tái hiện ($f_2 = m - 1$). Tra bảng ta có được giá trị $F_{tb} = F_{(0,95)(f_1, f_2)}$. So sánh F và F_{tb} , nếu $F < F_{tb}$ thì phương trình hồi quy tìm được là tương thích với thực nghiệm với mức ý nghĩa 95% , ngược lại là không tương thích.

Giá trị F tính theo công thức (10) với phương sai dư tính theo công thức (11): Các giá trị \hat{Y}_i tính được theo phương trình hồi qui (3): $\hat{Y}_i = \sum_{i=1}^4 b_i X_i$

Thay số vào phương trình hồi quy, tính toán ta được các giá trị Y_i .
 Từ đó có: $S_{dur}^2 = 9,8.10^{-6}$, $F = 6,24$.

So sánh F với $F_{(1-p)(f_1, f_2)}$ trong đó $P = 0,05$; $1-P = 0,95$ là mức ý nghĩa; f_1 là bậc tự do của phương sai tương thích ($f_1 = N - L = 3$); f_2 là bậc tự do của phương sai tái hiện ($f_2 = m - 1 = 2$). Tra bảng ta có được giá trị $F_{(0,95)(f_1, f_2)} = F(0,05)(3,2) = 18,3$.

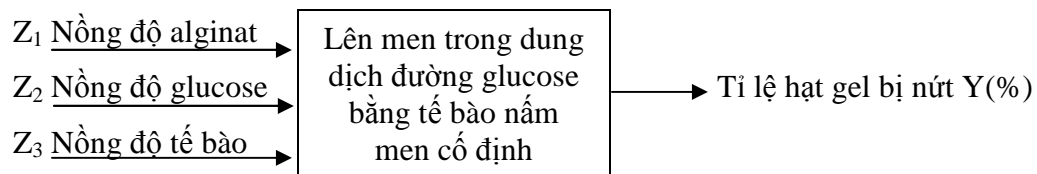
Như vậy $F < F_{tb}$, phương trình hồi quy tìm được là tương thích với thực nghiệm với mức ý nghĩa 95%.

5.5. Bài toán 5

➤ Mục đích: Nghiên cứu tối ưu hoá quy trình cố định tế bào nấm men bằng Alginat để lên men rượu.

➤ Quy trình công nghệ được mô tả theo sơ đồ (trang 1).

Nghiên cứu các yếu tố ảnh hưởng đến mạng lưới gel: nồng độ alginat; nồng độ glucose; nồng độ tế bào:



Sau quá trình lên men, vớt các hạt gel ra và xác định tỉ lệ (%) hạt gel bị nứt. Tỉ lệ hạt gel bị nứt càng thấp càng tốt nghĩa là hạt gel càng chắc càng tốt.

➤ Hàm mục tiêu: $Y = Y(Z_1, Z_2, Z_3)$

Bài toán tối ưu: Xác định nồng độ alginat; nồng độ glucose; nồng độ tế bào nấm men để hạt gel bền nhất trong quá trình lên men rượu bằng tế bào nấm men, cố định bằng alginat.

$$Y_{\min} = \min Y(Z_1, Z_2, Z_3)$$

Sau khi tiến hành các thí nghiệm thăm dò, tác giả đã chọn vùng khảo sát như sau:

$$Z_1 = 1 \div 4\%$$

$$Z_2 = 10 \div 18\%$$

$$Z_3 = 10 \div 20\%$$

Đây là bài toán tối ưu phỏng định, giải bài toán theo phương pháp leo dốc.

Phương án qui hoạch thực nghiệm: phương pháp trực giao cấp 1.

Số thí nghiệm phải làm: $2^k = 2^3 = 8$

$$\text{Với } Z_{1\min} = 1 \quad Z_1 \quad 4 = Z_{1\max}$$

$$Z_{2\min} = 10 \quad Z_2 \quad 18 = Z_{2\max}$$

$$Z_{3\min} = 10 \quad Z_3 \quad 20 = Z_{3\max}$$

Điểm xuất phát ở tâm phương án: $Z_0 = (2.5; 14; 15)$

Giá trị của hàm mục tiêu tại điểm Z_0 được xác định bằng thực nghiệm: $Y(Z_0) = 7.5$

Ma trận thực nghiệm được bố trí như sau:

Số TN	Z ₁	Z ₂	Z ₃	Y
1	4	18	20	12.35
2	4	18	10	8.87
3	4	10	20	12.08
4	4	10	10	6.92
5	1	18	20	42.13
6	1	18	10	13.51
7	1	10	20	22.19
8	1	10	10	4.57

Phương trình hồi qui có dạng:

$$Y = B_0 + B_1Z_1 + B_2Z_2 + B_3Z_3$$

Trong hệ mã hoá không thứ nguyên ta có được:

Mức trên: - kí hiệu +1

Mức cơ sở: - kí hiệu 0

Mức dưới: - kí hiệu -1

Công thức chuyển từ hệ đơn vị thực qua đơn vị mã hoá không thứ nguyên:

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

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

Thu được ma trận thực nghiệm với các biến mã như sau:

Số TN	X ₀	X ₁	X ₂	X ₃	Y
1	1	1	1	1	12.35
2	1	1	1	-1	8.87
3	1	1	-1	1	12.08
4	1	1	-1	-1	6.92
5	1	-1	1	1	42.13
6	1	-1	1	-1	13.51
7	1	-1	-1	1	22.19

8	1	-1	-1	-1	4.57
---	---	----	----	----	------

Từ kết quả thực nghiệm, tính toán các hệ số B_j :

$$B_0 = \frac{\sum_{i=1}^n Y_i}{N}$$

$$B_i = \frac{\sum_{i=1}^n X_{ij} Y_i}{N}$$

$$B_{j1} = \frac{\sum_{i=1}^n (X_j \cdot X_1)_i \cdot Y_i}{N}$$

Từ số liệu thực nghiệm trên, áp dụng các công thức trên ta xác định được giá trị B_0 , B_1 , B_2 và B_3 thu được kết quả:

B_1	B_2	B_3	B_0
-5.25	3.75	6.6	12.275

➤ Để tính phương sai tái hiện, tác giả làm thêm 3 thí nghiệm ở tâm.

Kết quả các thí nghiệm ở tâm:

N_0	Y_u^0	\bar{Y}^0	$Y_u^0 - \bar{Y}^0$	$(Y_u^0 - \bar{Y}^0)^2$	$\Sigma(Y_u^0 - \bar{Y}^0)^2$
1	5.65	7.50333	-1.8533	3.43472	8.22749
2	7.19		-0.3133	0.09818	
3	9.67		2.1667	4.69459	

Phương sai tái hiện được tính theo công thức:

$$S_{th}^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i^0 - \bar{Y}^0)^2$$

$$S_{th} = \sqrt{\frac{\sum_{i=1}^m (Y_i^0 - \bar{Y}^0)^2}{m-1}}$$

trong đó m là số thí nghiệm ở tâm phương án.

- Sự có nghĩa của hệ số hồi quy được kiểm định theo tiêu chuẩn Student:

$$t_i = \frac{|b_i|}{S_{b_i}}$$

b_i : là hệ số thứ i trong phương trình hồi quy.

S_{b_i} : Độ lệch quân phương của hệ số thứ i .

$$S_{b_i} = \frac{S_{th}}{\sqrt{N}}$$

➤ Phương sai tái hiện: $S_{th}^2 = 4.11$

Để kiểm định ý nghĩa các hệ số, tác giả tính các hệ số t_j , thu được kết quả sau:

t_0	t_1	t_2	t_3
23.3746	3.5263	5.42122	9.6644

Tra bảng phân phối phân vị Student với mức ý nghĩa $p = 0.05$, $f = N_0 - 1 = 2$ ta có $t_{0.05}(2) = 4.3$. Vậy các hệ số t_j đều lớn hơn $t_{0.05}(2)$ nên hệ số b_1 của phương trình hồi qui không có nghĩa.

Phương trình hồi qui có dạng sau:

$$\hat{Y}_L = 12.275 + 3.75X_2 + 6.6X_3$$

➤ Kiểm định sự tương thích của phương trình hồi qui với thực nghiệm:

STT	\hat{Y}_L	Y_i	$Y_i - \hat{Y}_L$	$(Y_i - \hat{Y}_L)^2$
1	20.81	12.35	-8.46	71.5716
2	7.09	8.87	1.78	3.1684
3	13.03	12.08	-0.95	0.9025
4	-0.69	6.92	7.61	57.9121
5	31.35	42.13	10.78	116.208
6	17.63	13.51	-4.12	16.9744
7	23.57	22.19	-1.38	1.9044
8	9.85	4.57	-5.28	27.8784

Phương sai dư (theo công thức trang 5): $S_{\text{dư}}^2 = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N - L}$

(N là số thí nghiệm, L là hệ số ý nghĩa)

Ta có: $S_{\text{d}}^2 = 74.13$

Tiêu chuẩn Fisher:

$$F = S_{\text{d}}^2 / S_{\text{th}}^2 = 74.13 / 4.1 = 18.08$$

Tra bảng phân vị phân bố Fisher với $p = 0.05$; $f_1 = N - L = 4$; $f_2 = N_0 - 1 = 2$; ta có:

$F_{1-p} = F_{0.095}(4,2) = 19.3$. Vậy $F < F_{0.95}(4;2)$. Phương trình hồi qui tương thích với thực nghiệm.

TÀI LIỆU THAM KHẢO

- [1] Nguyễn Thị Lan, *Quy hoạch thực nghiệm*, bài giảng dành cho sinh viên chuyên ngành Công nghệ Hóa học trường Đại học Bách khoa - Đại học Đà Nẵng.
- [2] Trần Văn Ngũ, *Lý thuyết thực nghiệm*, Đại học Bách Khoa TPHCM, 1997.
- [3] GS.TSKH Nguyễn Minh Tuyển, *Quy hoạch thực nghiệm*, Nhà XB Khoa học và Kỹ thuật, 2005.
- [4] GS.TSKH Nguyễn Minh Tuyển, PGS.TS Phạm Văn Thiêm, *Kỹ thuật hệ thống công nghệ hóa học*, Tập 1, Nhà XB Khoa học và Kỹ thuật, 2001.
- [5] X.L. Acnadarova, V.V. Capharov, *Tối ưu hóa thực nghiệm trong hóa học và công nghệ hóa học* (Tiếng Nga & bản dịch của Nguyễn Cảnh, Nguyễn Đình Soa), Đại học Bách Khoa TPHCM, 1994.
- [6] Živorad R. Lazić, **Design of Experiments in Chemical Engineering: A Practical Guide**, WILEY-VCH Verlag GmbH, 2004.
- [7] D.R. Cox and N. Reid, **The theory of design experiment**, Chapman & Hall/CRC, 2000.