

## บทที่ 3

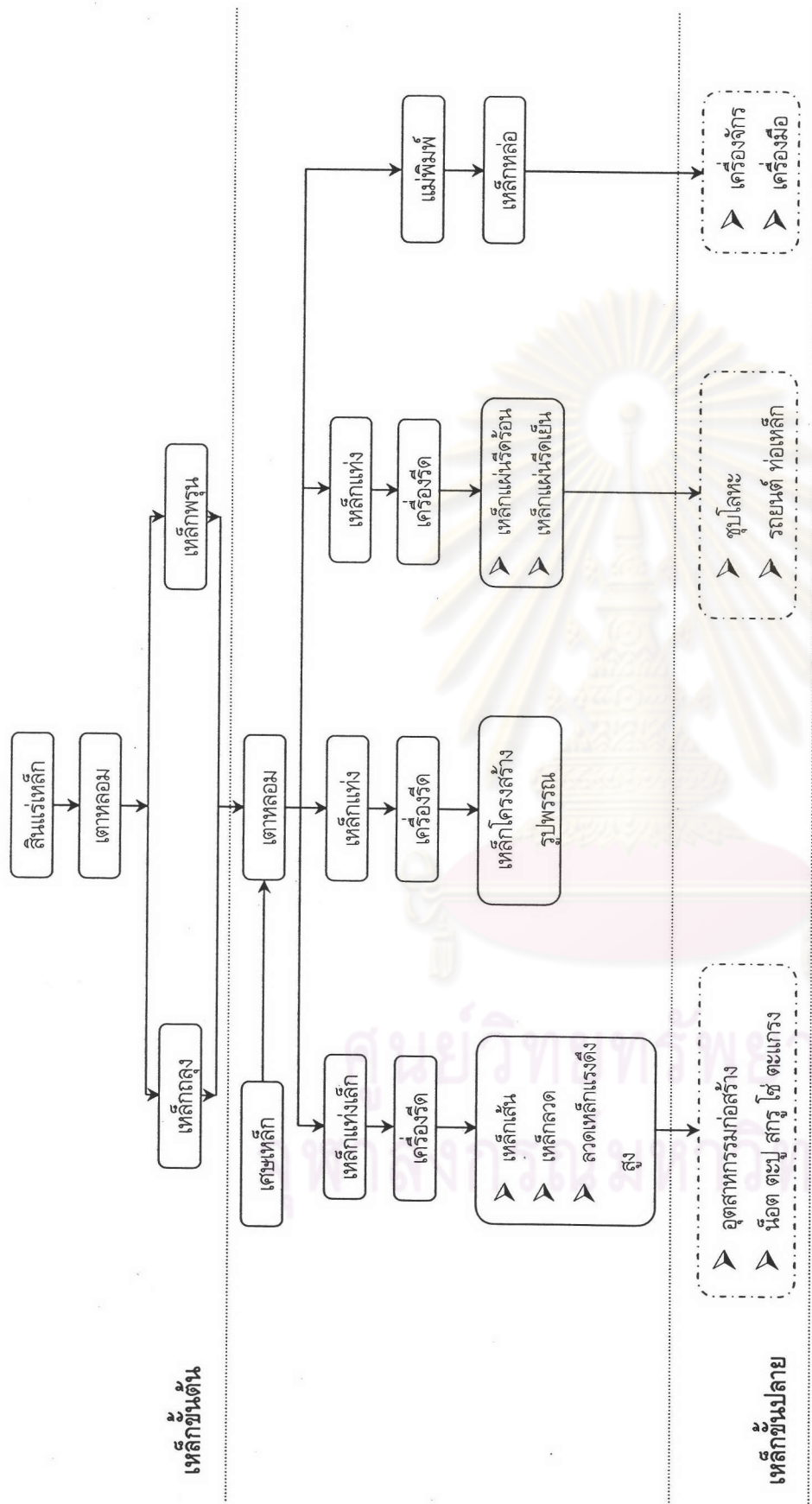
### ทฤษฎี

#### 3.1 กระบวนการแปรรูปเหล็ก และกระบวนการกำจัดสนิม

##### 3.1.1 กระบวนการแปรรูปเหล็ก

อุตสาหกรรมการผลิตเหล็กและเหล็กกล้า ประกอบด้วยผลิตภัณฑ์ชนิดต่างๆ มากมาย สามารถแบ่งเป็นประเภทใหญ่ๆ ตามขั้นตอนหรือกระบวนการผลิต ประกอบด้วย

1. เหล็กขั้นต้น (Raw Steel Product) ได้จากการนำสินแร่เหล็กมาเข้าเตาหลอมและ ทำการถลุงเหล็กหรือเป็นการขจัดออกซิเจนโดยใช้อากาศ (Direct Reduction) เหล็กที่ได้จากการถลุงเรียกเหล็กถลุง (Pig Iron) และเหล็กฟูน (Sponge Iron) ซึ่งใช้เป็นวัตถุดิบพื้นฐานในการผลิตเหล็ก (Steel Making)
2. เหล็กขั้นกลาง (Semi-Finished Steel Products) เป็นการนำเหล็กฟูน เหล็กถลุงและเศษเหล็ก (Scrap) มาหลอม ผลิตภัณฑ์เหล็กขั้นกลางที่ได้ คือ เหล็กแท่งกลม (Billet) เหล็กแท่งแบน (Slab) เหล็กแท่งใหญ่ (Bloom Beam)
  - เหล็กแท่งกลม (Billet) เป็นผลิตภัณฑ์ทรงกลมยาว เมื่อผ่านกระบวนการเหล็กขั้นปลาย ผลิตภัณฑ์ที่ได้คือ เหล็กเส้น, ลวดเหล็ก, ลวดเหล็กแรงดึงสูง เป็นต้น
  - เหล็กแท่งแบน (Slab) เป็นผลิตภัณฑ์ทรงแบนเรียบ เมื่อผ่านกระบวนการเหล็กขั้นปลาย ผลิตภัณฑ์ที่ได้คือ เหล็กแผ่นรีดร้อนและรีดเย็น
  - เหล็กแท่งใหญ่ (Bloom, Beam) เมื่อผ่านกระบวนการเหล็กขั้นปลาย ผลิตภัณฑ์ที่ได้คือ เหล็กโครงสร้างรูปพรรณ เช่น เหล็กคาน, เหล็กฉาก, ตัว I, T, H,
  - เศษเหล็ก (Scrap) ส่วนใหญ่จะได้เป็นผลิตภัณฑ์เหล็กหลอมขึ้นรูป
3. เหล็กขั้นปลาย (Finish Steel Products) เป็นการนำผลิตภัณฑ์ขั้นกลางมาผ่านกระบวนการแปรรูป ทำได้ 2 ขั้นตอน คือ การแปรรูปรีดร้อนและการแปรรูปรีดเย็น ได้ผลิตภัณฑ์สำเร็จรูป ได้แก่
  - ผลิตภัณฑ์เหล็กทรงยาว (Long Product) เช่น เหล็กเส้น, เหล็กข้ออ้อย, ลวดที่เป็นโครงสร้างรูปพรรณ เช่น เหล็กคาน, เหล็กฉาก, ตัว I, T, H เป็นต้น
  - ผลิตภัณฑ์เหล็กทรงแบน (Flat Product) จะได้ผลิตภัณฑ์เหล็กรีดร้อน สามารถใช้ในอุตสาหกรรมยานยนต์ ท่อเหล็ก ก่อสร้าง, ผลิตภัณฑ์เหล็กรีดเย็น ใช้ในอุตสาหกรรมไฟฟ้า บรรจุภัณฑ์ เป็นต้น และเหล็กแผ่นเคลือบผิว
  - ผลิตภัณฑ์เหล็กรูปพรรณ (Shape/Section) ผลิตภัณฑ์ที่ผ่านกรรมวิธีรีดร้อน, รีดเย็นทำเป็นรูปพรรณ เช่น เหล็กคาน เหล็กฉาก, ตัว I, T, H เป็นต้น



แหล่งข้อมูล : ธนาคารแห่งประเทศไทย

= กลุ่มอุตสาหกรรมที่เกี่ยวข้องกับกระบวนการ

รูปที่ 3.1 แสดงโครงสร้างอุตสาหกรรมการแปรรูปเหล็กในประเทศไทย

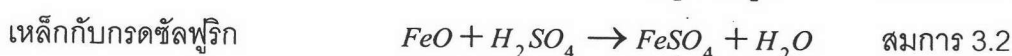
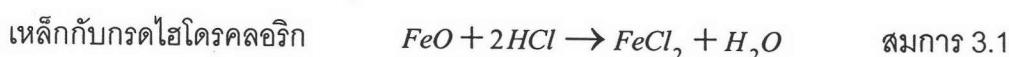
### 3.1.2 กระบวนการกำจัดสนิม

กระบวนการกำจัดสนิม เป็นการกำจัดคราบออกไซด์ออกก่อนที่จะมีการแปรรูปหรือการดำเนินการในขั้นตอนต่อไป เช่น การชุบสังกะสี, การเคลือบ, การทาสี, หรือการชุบด้วยไฟฟ้า เป็นต้น ประสิทธิภาพในการกำจัดสนิมขึ้นอยู่กับตัวแปรหลายตัว เช่น เวลาในการกำจัดสนิม ความเข้มข้นของสารละลายกรดที่ใช้ และอุณหภูมิของสารละลายกรด เวลาที่ใช้ในการกำจัดสนิมควรจะสั้นในกรณีที่ปริมาณของสนิมมีน้อย และการผลิตมีจำนวนมาก ความเข้มข้นของสารละลายกรดต้องเหมาะสมกับเวลา ถ้าความเข้มข้นของสารละลายกรดมีสูงไม่ควรใช้เวลาในการกำจัดสนิมนานเกินไปเพื่อหลีกเลี่ยงการกำจัดสนิมที่มากเกินไปจนถึงเนื้อเหล็ก สำหรับอุณหภูมิของสารละลายกรดไม่ควรสูงเกินไปเพื่อป้องกันการเกิดไอกรดในปริมาณที่สูง

ขั้นตอนในการกำจัดสนิมแบ่งออกเป็น 2 ขั้นตอนด้วยกัน คือ ขั้นตอนการทำความสะอาดเบื้องต้น และขั้นตอนการกำจัดสนิม ดังแสดงในรูปที่ 3.2

3.1.2.1. ขั้นตอนการทำความสะอาดเบื้องต้น เป็นการล้างวัตถุบดด้วยน้ำสะอาด แต่ในบางกรณีจะมีการใช้สารเคมี เช่น โซเดียมไฮดรอกไซด์ อัลคาไลน์ซิลิเกต และฟอสเฟต ในน้ำล้างเพื่อเพิ่มประสิทธิภาพในการทำทำความสะอาด การทำความสะอาดเบื้องต้นที่นิยมจะกระทำโดยการจุ่มวัตถุบดลงในอ่าง หรือพ่นด้วยสารละลายน้ำล้างลงบนวัตถุบด เมื่อฝุ่นและน้ำมันถูกกำจัดออกจากผิวโลหะแล้ว วัตถุบดจะถูกส่งไปยังขั้นตอนการกำจัดสนิม ซึ่งถ้าหากมีน้ำมันติดอยู่บนผิวโลหะ การกำจัดสนิมด้วยกรดจะทำได้ไม่ดี ดังนั้นขั้นตอนการทำความสะอาดเบื้องต้นจะต้องแน่ใจว่ากำจัดฝุ่นและน้ำมันออกจากผิวโลหะหมดแล้ว

3.1.2.2. ขั้นตอนการกำจัดสนิม ชิ้นงานจะถูกกำจัดสนิมโดยการนำไปจุ่มในอ่างที่มีสารละลายกรดไฮโดรคลอริกหรือกรดซัลฟูริกอยู่ในระยะเวลาหนึ่งที่เพียงพอที่จะกำจัดคราบออกไซด์ออกหมด ซึ่งกรดทั้งสองจะมีข้อดีข้อเสียต่างกัน เช่นกรดไฮโดรคลอริกเป็นกรดที่มีความสามารถในการกัดกร่อนสูง แต่มีปัญหาเกี่ยวกับปริมาณไอกรดที่เกิดขึ้นจากกระบวนการ โดยเฉพาะอย่างยิ่งที่ความเข้มข้นและอุณหภูมิสูง สำหรับกรดซัลฟูริกนั้นจะมีการกัดกร่อนไม่ดีเท่ากับกรดไฮโดรคลอริก แต่ปัญหาเรื่องไอกรดก็มีน้อยตามไปด้วย อย่างไรก็ตามในประเทศไทยพบว่ากรดไฮโดรคลอริกจะเป็นที่นิยมมากกว่า ปฏิกริยาระหว่างออกไซด์ของเหล็กและกรดสามารถแสดงได้ดังนี้





ขั้นตอนการกำจัดสนิมประกอบด้วยขั้นตอนการกำจัดสนิมและขั้นตอนการล้าง ขั้นตอนการกำจัดสนิมคือ ขั้นตอนการกำจัดสนิมบนผิวโลหะด้วยสารละลายกรดเพื่อทำให้ผิวเหล็กสะอาด จากนั้นสารละลายกรดที่ติดไปกับชิ้นงานจะถูกกำจัดออกจากผิวเหล็กโดยการล้างด้วยน้ำใน ขั้นตอนการล้างก่อนที่เหล็กจะถูกส่งไปยังกระบวนการอื่นต่อไป เป็นที่สังเกตว่าอ่างกำจัดสนิมและอ่างล้างที่ใช้โดยทั่วไปในขั้นตอนการกำจัดสนิมและขั้นตอนการล้างตามลำดับ อาจจะประกอบด้วยชุดของอ่างต่อแบบอนุกรมกันในแต่ละขั้นตอน นอกจากนี้ยังมีการเติมน้ำเข้าไปในอ่างล้างนี้เพื่อเพิ่มอุณหภูมิของสารละลายกรดและน้ำล้าง ซึ่งจะเป็นการเพิ่มประสิทธิภาพของการกำจัดสนิมและการทำความสะอาดด้วย

สำหรับกระบวนการกำจัดสนิมสามารถแบ่งออกเป็นกระบวนการแบบกะ (Batch) และกระบวนการแบบต่อเนื่อง (Continuous) สำหรับกระบวนการกำจัดสนิมแบบกะ เหล็กจะถูกเคลื่อนยกขึ้น เพื่อนำไปจุ่มที่อ่างต่างๆ เมื่อมาถึงอ่างกำจัดสนิมหลังจากเสร็จสิ้นการกำจัดสนิมแล้วเคลื่อน จะทำการยกเหล็กขึ้นแล้วทิ้งไว้ระยะหนึ่งให้กรดที่เกาะอยู่บนผิวชิ้นงานไหลกลับลงสู่อ่าง เพื่อที่จะลดการสูญเสียสารละลายกรดที่ใช้ นอกจากนี้ จะมีแผ่นลาดเอียงวางพาดอยู่ระหว่างอ่างเพื่อทำให้สารละลายกรดที่หยดระหว่างการยกนั้นไหลกลับไปยังอ่างก่อนหน้านี้ กระบวนการแบบกะจะเป็นที่นิยมสำหรับวัสดุที่เป็นรูปทรงอื่นๆ เช่น ลวด ท่อ สกรู สำหรับกระบวนการกำจัดสนิมแบบต่อเนื่อง จะมีชุดลูกกลิ้งไว้สำหรับเคลื่อนย้ายเหล็กไปที่บ่อต่างๆ ช่องห่างระหว่างลูกกลิ้งที่เหมาะสมจะช่วยลดปริมาณสารละลายกรดที่ติดไปกับแผ่นเหล็กได้ โดยทั่วไปกระบวนการแบบต่อเนื่องจะเป็นที่นิยมสำหรับวัสดุที่เป็นแผ่นเรียบต่อเนื่อง

### 3.1.3 แหล่งกำเนิดของเสีย

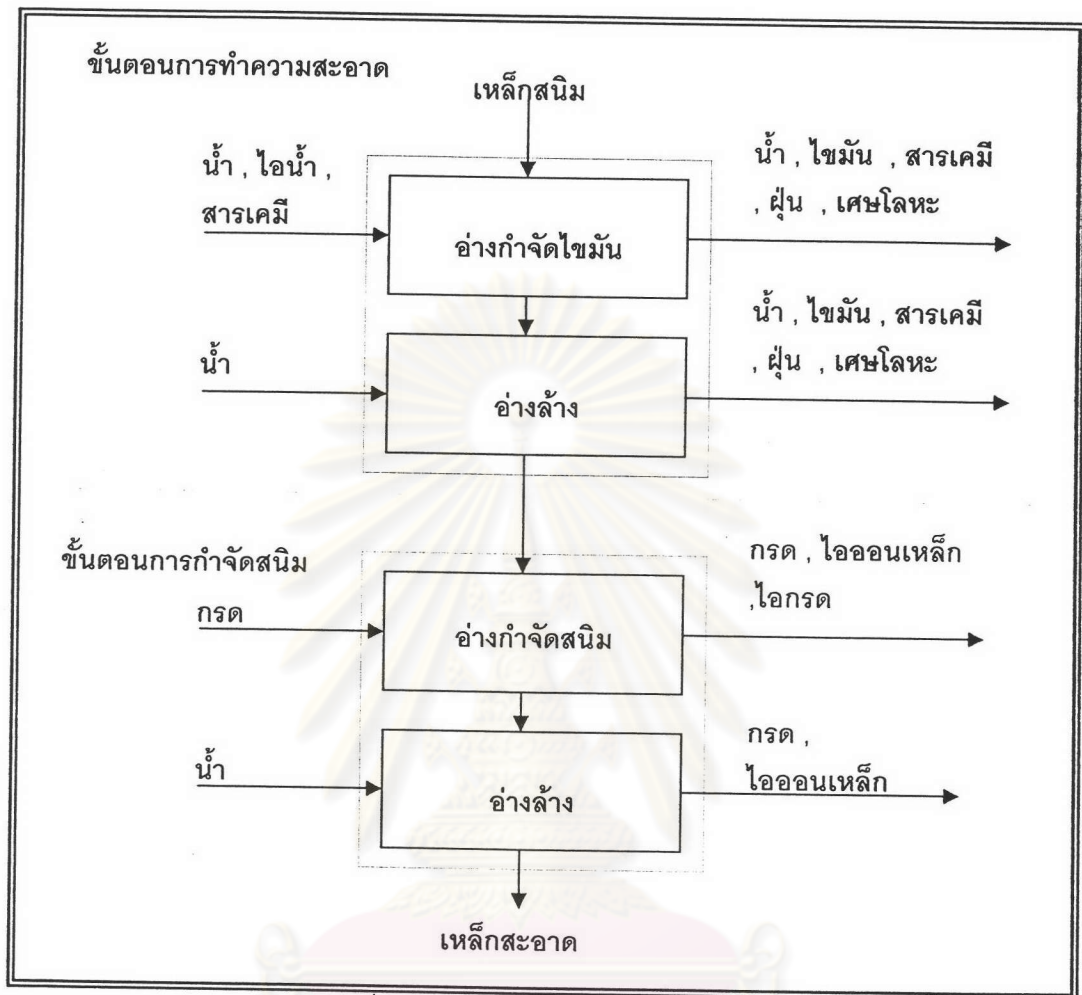
#### 3.1.3.1. ข้อมูลทั่วไป

กระบวนการกำจัดสนิมประกอบด้วยขั้นตอนการทำความสะอาดเบื้องต้นและขั้นตอนการกำจัดสนิม ซึ่งในขั้นตอนการกำจัดสนิมจะมีขั้นตอนย่อย คือ ขั้นตอนของการกำจัดสนิมและขั้นตอนการล้าง ดังนั้นแหล่งกำเนิดของเสียของกระบวนการกำจัดสนิมสามารถแบ่งออกได้เป็นอ่างทำความสะอาดเบื้องต้น อ่างกำจัดสนิม และอ่างล้าง

เนื่องจากกระบวนการกำจัดสนิมเป็นกระบวนการกำจัดสนิมและฟิล์มออกไซด์ออกจากผิวโลหะโดยใช้สารละลายกรดและด่างเพื่อที่จะได้ผิวชิ้นงานที่ไม่มีออกไซด์ ของเสียหลักของกระบวนการนี้คือ น้ำเสียที่มาจากน้ำล้างและสารละลายที่ใช้กำจัดสนิม อย่างไรก็ตามของเสียยังปรากฏอยู่ในรูปของแข็งและก๊าซ ของเสียที่อยู่ในรูปของของแข็งนั้นมาจากฝุ่นบนผิวโลหะและเศษโลหะเป็นต้น ฝุ่นและสนิมจะสะสมที่ผิวโลหะระหว่างการขนส่งและการเก็บรักษาวัสดุ สำหรับของเสียที่อยู่ในรูปของก๊าซนั้นส่วนใหญ่จะเป็นไฮโดรเจนจากอ่างกำจัดสนิมหรือก๊าซที่เกิดขึ้นจาก



ปฏิกิริยาในอ่างกำจัดสนิม รูปที่ 3.2 เป็นการสรุปชนิดของของเสีย (ตามสถานะ) ที่เกิดขึ้นจากอ่าง ทำความสะอาดเบื้องต้น อ่างกำจัดสนิมและอ่างล้าง



รูปที่ 3.2 แสดงกระบวนการกำจัดสนิม

### 3.1.3.2. แหล่งของเสีย

#### 3.1.3.2.1 อ่างทำความสะอาดเบื้องต้น

เหล็กและวัตถุที่เป็นเหล็กจะถูกทำความสะอาดเบื้องต้นในอ่างทำความสะอาดเบื้องต้นโดยใช้น้ำเปล่าในการล้างเพื่อกำจัดฝุ่นและน้ำมันหรือจารบีจากผิวโลหะ สารเคมี เช่น โซเดียมไฮดรอกไซด์ อัลคาไลน์ซิลิเกต และฟอสเฟต อาจถูกเติมเข้าไปในน้ำล้างเพื่อเพิ่มประสิทธิภาพในการล้าง นอกจากนี้การเพิ่มประสิทธิภาพในการล้างยังสามารถทำได้โดยการเพิ่มอุณหภูมิของน้ำล้าง การทำความสะอาดเบื้องต้นสามารถกระทำโดยการจุ่มเหล็กและวัตถุที่เป็นเหล็กลงในน้ำล้างในอ่างทำความสะอาดเบื้องต้น โดยการพ่นน้ำที่ผิว หรือทั้งสองวิธีรวมกัน ของเสียที่เกิดขึ้นจากการทำความสะอาดเบื้องต้นที่อ่างทำความสะอาดเบื้องต้นคือน้ำเสียที่ประกอบด้วยฝุ่น น้ำมันและจารบีและสารเคมีต่าง ๆ ที่เติมลงไป

### 3.1.3.2.2 อ่างกำจัดสนิม

ของเสียส่วนใหญ่จากกระบวนการกำจัดสนิมจะเกิดขึ้นในอ่างกำจัดสนิมน้ำเสียของกระบวนการนี้ คือ สารละลายกรดซึ่งประกอบด้วย กรด ดิวบยั้ง ไอออนเหล็ก และสารเคมีที่ใช้ในอ่างทำความสะอาดเบื้องต้น กระบวนการกำจัดสนิมเกือบทั้งหมดในประเทศไทยใช้กรดไฮโดรคลอริก เป็นสารละลายกรดสำหรับกำจัดสนิม อย่างไรก็ตาม บางโรงงานมีการใช้กรดซัลฟูริกแทนและสารเคมีเหล่านี้จึงเป็นองค์ประกอบที่อยู่ในน้ำเสียจากอ่างกำจัดสนิม

### 3.1.3.2.3 อ่างล้าง

มีไว้เพื่อกำจัดสารละลายที่ใช้ในขั้นตอนการกำจัดสนิมที่ติดมากับผิวโลหะออก การล้างนี้ทำโดยการจุ่มเหล็กหรือชิ้นงานลงในน้ำในอ่างน้ำ โดยฉีดน้ำบนผิวโลหะ การใช้ทั้ง 2 วิธี โดยทั่วไปจะไม่มีสารเติมสารเคมีใด ๆ ลงในอ่างล้าง ดังนั้นสารประกอบหลักของน้ำเสียจากอ่างล้างจะเหมือนกับของเสียจากอ่างกำจัดสนิมแต่ความเข้มข้นจะน้อยกว่ามาก ส่วนมากของเสียสามารถพบได้ในอ่างล้างเช่นกันแต่ในปริมาณน้อยมาก

## 3.2 การแลกเปลี่ยนไอออน และ เรซินแลกเปลี่ยนไอออน

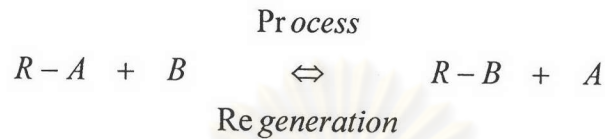
จากที่ได้ศึกษาเรื่องการดูดซับ เราได้ว่าการแลกเปลี่ยนไอออนเป็นวิธีการแยกหนึ่งของการดูดซับโดยมีลักษณะที่แตกต่างกันคือ การดูดซับปกติภายในรูพรุนจะมีช่องว่างจำนวนมาก สำหรับการดูดซับสารที่ต้องการดูดซับจากสารละลายหนึ่ง ๆ ส่วนการแลกเปลี่ยนไอออนภายในรูพรุนจะเต็มไปด้วยไอออนที่พร้อมจะถ่ายเทและแลกเปลี่ยนกับ ไอออนที่ต้องการดูดซับโดยลักษณะการทำงานของเรซินแลกเปลี่ยนไอออนเป็นดังนี้

### 3.2.1 การแลกเปลี่ยนไอออน

การแลกเปลี่ยนไอออน (Ion exchange) เป็นการแลกเปลี่ยนระหว่างไอออนที่อยู่ในภาคของของแข็งกับไอออนของของเหลวที่ไม่สามารถละลายซึ่งกันและกันได้ และทั้งสองส่วนนี้ต้องอยู่ติดกัน ไอออนที่อยู่ในภาคของแข็งเป็นประจุที่ไม่ติดแน่นหรือมีโครงสร้างแบบเปิด กล่าวคือสามารถยอมให้ไอออนและโมเลกุลของตัวทำละลายสามารถเคลื่อนที่ได้อย่างอิสระทั้งเข้าและออกซึ่งเกิดจากแรง อิเล็กโตรสแตติก (Electrostatic force) บนผิวของเรซิน (Resin) และเกิดการแลกเปลี่ยนไอออนที่มีประจุเหมือนกันในสารละลายที่ไหลผ่านเรซิน โดยอาศัยหลักการที่ไอออนแต่ละชนิดมีความชอบหรือถูกดูดซับโดยเรซินไม่เท่ากัน หรือเป็นการแลกเปลี่ยนไอออนระหว่างของแข็งกับของเหลว โดยไม่เกี่ยวข้องกับอนุโมลในโครงสร้างทางกายภาพของของแข็งแต่อย่างใด ไอออนที่อยู่ในภาคของแข็งจะต้องมีปริมาณที่มากพอเพื่อให้มีการแลกเปลี่ยนไอออนได้อย่าง



รวดเร็ว การแลกเปลี่ยนไอออนมีทั้งการแลกเปลี่ยนไอออนลบ (Anion) และการแลกเปลี่ยนไอออนบวก (Cation) เมื่อเกิดการแลกเปลี่ยนแล้วสามารถเปลี่ยนกลับมาอยู่ในรูปของไอออนเดิม โดยการผ่านสารละลายที่มีไอออนเช่นเดียวกับไอออนเริ่มต้น เรียกขั้นตอนนี้ว่า การรีเจนเนอเรต (Regeneration) และสารที่ใช้เรียกว่า รีเจนเนอเรนท์ (Regenerant) การแลกเปลี่ยนไอออนแสดงดังสมการ

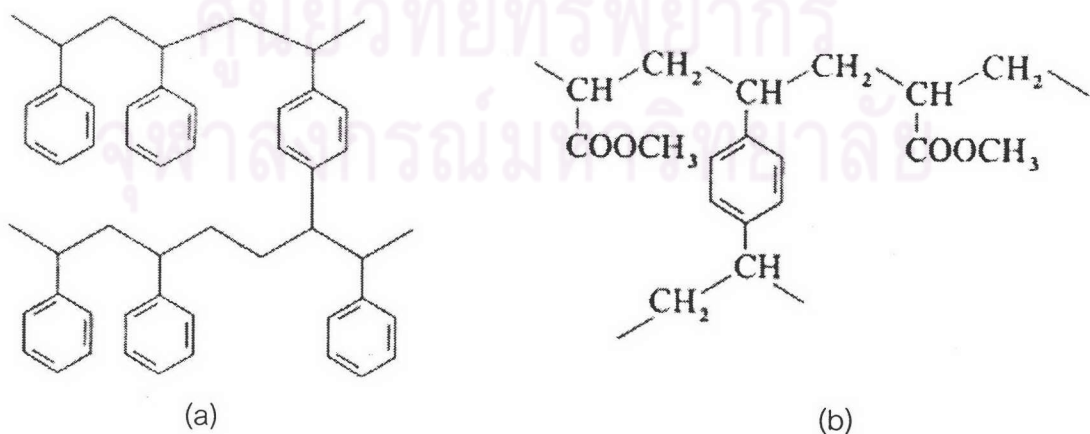


อัตราการแลกเปลี่ยนไอออน ระหว่างไอออนบนผิวของเรซินกับไอออนในสารละลายถูกควบคุมด้วยอัตราการไหลของสารละลาย ที่อัตราการไหลช้าจะเกิดการแลกเปลี่ยนไอออนได้ดีกว่า อัตราการไหลเร็ว และการแลกเปลี่ยนไอออนยังถูกนำมาประยุกต์ใช้ในงานสำหรับปรับปรุงคุณภาพน้ำ และของเสียอื่นๆ เช่น การแทนที่ หรือการแยกไอออนของโลหะที่มีค่า ( $\text{Cu}^{2+}$ ) หรือโลหะมีพิษ (ไซยาไนด์) ออกจากของเสียก่อนทิ้งลงสู่ลำธารสาธารณะ และการกำจัด หรือลดความกระด้างของน้ำ (Softening) ซึ่งเป็นการแลกเปลี่ยนไอออนของแคลเซียม ( $\text{Ca}^{2+}$ ) และแมกนีเซียม ( $\text{Mg}^{2+}$ ) ที่ทำให้น้ำกระด้าง กับไอออนของโซเดียม ( $\text{Na}^+$ ) ของเรซิน และการกำจัดไอออนต่างๆ ที่อยู่ในน้ำ (Deionization) เพื่อให้ได้น้ำที่มีความบริสุทธิ์สูง เนื่องจากมีต้นทุนการผลิตที่ต่ำกว่าวิธีอื่นมาก อีกทั้งยังสิ้นเปลืองพลังงานน้อยกว่าวิธีอื่นๆ อีกด้วย

โครงสร้างของเรซินแลกเปลี่ยนไอออนประกอบด้วยโมเลกุล 2 ส่วน ส่วนแรกคือส่วนโพลิเมอร์ (Polymer Matrices) ที่เรียกว่า เรซิน (Resin) โดยธรรมชาติมีโครงสร้างที่ซับซ้อน สามารถหาได้จากธรรมชาติ และสังเคราะห์ขึ้นและโครงสร้างเหล่านี้เกิดจากการโพลิเมอไรเซชัน (Polymerization) ของสไตรีน (Styrene) หรืออะครีเลต (Acrylate) เป็นต้น ซึ่งการโพลิเมอไรเซชันจะได้โมเลกุลที่มีลักษณะเป็นเส้นตรง โดยทั่วไปจะมีไดไวนิลเบนซีน (Divinylbenzene, DVB) เป็นตัวเชื่อมระหว่างสารที่ทำการโพลิเมอไรเซชันหลายๆสาย ซึ่ง DVB นี้เป็นตัวกำหนดความแข็งแรงของโครงสร้างเรซิน ส่วนที่สองคือหมู่ฟังก์ชันของไอออน (Functional Groups) ติดอยู่ และถูกทำให้สมดุลด้วยไอออนที่มีประจุตรงข้าม ซึ่งไอออนที่มีประจุตรงข้ามนี้ จะใช้เป็นไอออนที่ทำการแลกเปลี่ยนกับไอออนที่อยู่ในสารละลาย โดยในปี ค.ศ. 1945 D'Allelio ได้จดสิทธิบัตรการสังเคราะห์สารอินทรีย์ ที่เกิดจากการทำปฏิกิริยาโพลิเมอไรเซชัน (Polymerization) ระหว่างสารสไตรีน (Styrene) หรือ อะครีเลต (Acrylate) กับไดไวนิลเบนซีน (Divinylbenzene, DVB) มีโครงสร้างดังรูปที่ 3.3(a) โดยมีลักษณะเป็นของแข็งที่มีการเชื่อมต่อระหว่างโมเลกุล โดยโมเลกุลของสไตรีนเป็นโครงสร้างพื้นฐานและมีไดไวนิลเบนซีน เชื่อมต่อแบบไขว้สามมิติ (Three

dimensions cross linking) ทำให้สารโพลิเมอร์ที่ได้ไม่ละลายในตัวทำละลายใดๆ และมีความเหนียว อัตราส่วนระหว่างสไตรีนต่อไดไวนิลเบนซีนอยู่ระหว่าง ร้อยละ 1-20 โดยน้ำหนัก มีผลต่อโครงสร้างของเรซิน อัตราการเชื่อมโยงแบบสามมิติ (Degree of cross linking) ความพรุนของเรซิน และคุณสมบัติเฉพาะตัวของเรซิน ทำให้เรซินแต่ละชนิดมีคุณสมบัติทางด้านกายภาพและทางด้านเคมีแตกต่างกัน เช่น ความหนาแน่น (Density) การดูดซึมน้ำ (Hydration) การพองตัว (Swelling) การซึมผ่าน (Diffusion) และรูพรุนของเรซิน (Relative porosity) โดยทั่วไปจะให้อัตราส่วนของสไตรีนต่อไดไวนิลเบนซีนประมาณ ร้อยละ 8 เรซินที่สังเคราะห์ได้มีลักษณะเป็นเจล (Gel) การแลกเปลี่ยนไอออนเกิดจากการที่ไอออนซึมผ่านช่องของโครงสร้างเรซิน และเกิดการแลกเปลี่ยนกับไอออนที่เกาะอยู่บนเรซิน เรซินชนิดนี้ไม่สามารถใช้ได้กับสารที่มีขนาดโมเลกุลใหญ่กว่า 40 อังสตรอม ( $^{\circ}\text{A}$ ) ต่อมาในปี 1959 ได้มีการสังเคราะห์เรซินที่มีโครงสร้างแบบแมคโคพอร์รัส (Macroporous) โดยลดขนาดของช่องว่างภายในเล็กน้อย แต่มีขนาดของรูพรุนใหญ่ขึ้นประมาณ 1300 อังสตรอม โดยการเปลี่ยนแปลงอัตราส่วนผสมของสไตรีนกับไดไวนิลเบนซีน นอกจากนี้ยังทำให้เรซินมีความทนทานต่อการเปลี่ยนแปลงของอุณหภูมิ ทนต่อการเกิดออกซิเดชันได้มากขึ้น และมีอัตราการเปลี่ยนแปลงไอออนเร็วขึ้น

สารโพลิเมอร์ที่ใช้เป็นสารตั้งต้นเพื่อสังเคราะห์เรซินสำหรับแลกเปลี่ยนไอออนทั้งชนิด แคทไอออนเรซิน และแอนไอออนเรซิน โดยนำสารโพลิเมอร์มาทำปฏิกิริยากับสารเคมีกรดหรือด่าง เพื่อให้ได้หมู่ฟังก์ชันที่เกาะอยู่บนโครงสร้างของเรซินตามต้องการ เช่น การทำปฏิกิริยาซัลโฟเนชัน (Sulfonation) ดังในรูปที่ 3.3(b) จะได้เรซินที่มีคุณสมบัติการแลกเปลี่ยนประจุบวกแบบกรดแก่ (Strong acid cation exchanger) ซึ่งมีหมู่ฟังก์ชันของซัลโฟนิก ( $\text{HSO}_3$ ) เกาะติดกับโครงสร้างของเรซิน ไอออนของไฮโดรเจน ( $\text{H}^+$ ) ในหมู่ฟังก์ชันของซัลโฟนิกจะเป็นไอออนที่เกิดจากการแลกเปลี่ยนกับแคทไอออนที่อยู่ในสารละลาย

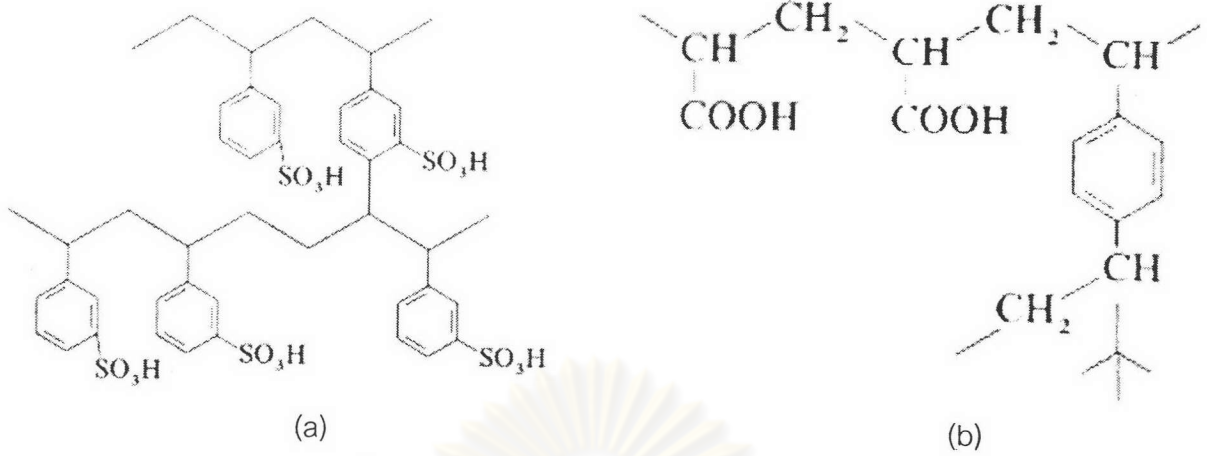


รูปที่ 3.3 แสดงโครงสร้างเรซินจากการโพลิเมอไรซ์ระหว่างสไตรีน (a) อะครีเลต (b) ไดไวนิลเบนซีน



ความสามารถของเรซินในการแลกเปลี่ยนไอออนในสารละลายขึ้นอยู่กับชนิดของไอออนและจำนวนของหมู่ฟังก์ชันที่เกาะอยู่บนเรซินนั้น ชนิดของไอออนที่เกาะบนเรซินที่มีผลต่อความสมดุล (Equilibrium) และความเฉพาะเจาะจง (Selectivity) ต่อเรซินมี 2 ชนิด คือชนิดแลกเปลี่ยนไอออนบวก (Cation Exchanger) และชนิดแลกเปลี่ยนไอออนลบ (Anion Exchanger) และเมื่อแบ่งตามหมู่ฟังก์ชัน สามารถแบ่งได้ 4 ชนิด คือ

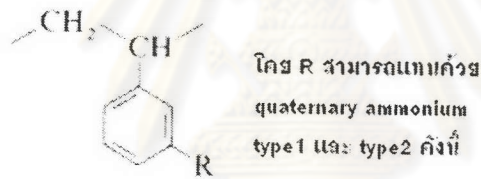
1. เรซินแลกเปลี่ยนไอออนบวกแบบกรดแก่ (Strong Acid Cation Exchanger, SAC) ได้แก่เรซินที่มีหมู่ฟังก์ชันซัลโฟนิก ( $\text{SO}_3\text{H}$ ) ฟอสโฟนิก ( $\text{H}_2\text{PO}_4$ ) และฟีนอลิก ( $\text{OH}$ ) แสดงในรูปที่ 3.4 (a)
2. เรซินแลกเปลี่ยนไอออนบวกแบบกรดอ่อน (Weak Acid Cation Exchanger, WAC) ได้แก่เรซินที่มีหมู่ฟังก์ชันคาร์บอกซิล (Carboxyl,  $\text{COO}^-$ ) แสดงในรูปที่ 3.4(b) เกิดจากการทำปฏิกิริยาระหว่างสารโพลีเมอร์ของสไตรีนกับไดไวนิลเบนซีน กับกรดคาร์บอนิก ( $\text{R-COOH}$ ) เรซินแบบ WAC มีคุณสมบัติกรดอ่อน ความสามารถในการแลกเปลี่ยนไอออนของไฮโดรเจนต่ำกว่าแบบ SAC แต่มีข้อดีคือสามารถรีเจนเนอเรต WAC ได้ง่ายกว่าแบบ SAC
3. เรซินแลกเปลี่ยนไอออนลบแบบต่างแก่ (Strong Basic Anion Exchanger, SBA) ได้แก่เรซินที่มีหมู่ฟังก์ชันของเอมีน (Amine) แสดงในรูปที่ 3.5(type1) ซึ่งได้จากการทำปฏิกิริยาเมทิลเลชัน (Methylation) ของโพลีเมอร์กับสารไตรเมทิลลามีน (Trimethylamine) หรือไดเมทิลเอทานอลามีน (Dimethyl ethanolamine)
4. เรซินแลกเปลี่ยนไอออนลบแบบต่างอ่อน (Weak Basic Anion Exchanger, WBA) ได้แก่เรซินที่มีหมู่ฟังก์ชันของเอมีนแบบเซคันดารีเอมีน (Secondary Amine,  $\text{CH}_2\text{NH}_2(\text{CH}_3)^+$ ) หรือเทอร์ติเอรีเอมีน (Tertiary Amine,  $\text{CH}_2\text{NH}(\text{CH}_3)_2^+$ ) แสดงในรูปที่ 3.5(type2) WBA มีคุณสมบัติคล้ายกับ WAC แบบตรงข้ามกัน สามารถรีเจนเนอเรตได้ด้วยโซเดียมไฮดรอกไซด์ แอมโมเนียมไฮดรอกไซด์ หรือโซเดียมคาร์บอเนต เรซินที่ใช้ในการแลกเปลี่ยนไอออนลบเป็นโพลีเมอร์ที่มีน้ำหนักโมเลกุลสูง



รูปที่ 3.4 แสดงตัวอย่างเรซินแลกเปลี่ยนประจุบวกโดย

(a) แสดงถึงเรซินแลกเปลี่ยนประจุลบแบบเบสแก่ (SAC)

(b) แสดงถึงเรซินแลกเปลี่ยนประจุลบแบบเบสอ่อน (WAC)



$-\text{CH}_2\text{N}^+(\text{CH}_3)_3\text{Cl}^-$  (type 1 resin)  
benzyltrimethylammonium groups

$-\text{CH}_2\text{N}^+(\text{CH}_3)_2\text{CH}_2\text{CH}_2\text{OHCl}^-$  (type 2 resin)  
benzyl dimethylethanolammonium groups

รูปที่ 3.5 แสดงตัวอย่างเรซินแลกเปลี่ยนประจุลบ โดย  
type1 แสดงถึงเรซินแลกเปลี่ยนประจุลบแบบเบสแก่ (SBA)

type2 แสดงถึงเรซินแลกเปลี่ยนประจุลบแบบเบสอ่อน (WBA)



### 3.2.2 ขั้นตอนการทำงานของเครื่องแลกเปลี่ยนไอออน

1. การเริ่มต้นทำงาน (Service cycle) หรือ Exhaustion เป็นการนำน้ำเสียผ่านเรซิน เพื่อให้เกิดการแลกเปลี่ยนไอออน น้ำที่ผ่านออกมาจะเป็นน้ำที่ปราศจากไอออน
2. การล้างย้อนกลับ (Backwash) เป็นการปล่อยน้ำให้ไหลสวนทางกับการผ่านน้ำเสีย เพื่อป้องกันเรซินอัดตัวแน่น และขจัดสิ่งเจือปนที่อาจติดอยู่ด้านบนของคอลัมน์
3. การรีเจนเนอเรต (Regeneration) เป็นการชะ (Elute) ไอออนที่เรซินจับไว้ออกมาด้วยสารละลายรีเจนเนอแรนต์ และทำให้เรซินกลับสู่สภาพเดิมเพื่อการใช้งานในรอบต่อไป
4. การล้างด้วยน้ำแบบช้า (Slow displacement rinse) เป็นขั้นตอนการล้างด้วยน้ำเพื่อกำจัดสารรีเจนเนอแรนต์ที่ตกค้างอยู่ออกให้หมด โดยใช้อัตราการไหลเดียวกับอัตราการไหลของการรีเจนเนอเรต
5. การล้างด้วยน้ำแบบเร็ว (Fast rinse) เป็นการล้างเพื่อเตรียมเรซินให้อยู่ในสภาพการใช้งานในรอบต่อไป โดยล้างด้วยอัตราการไหลเดียวกับอัตราการไหลของการผ่านน้ำเสียนานประมาณ 10 – 30 นาที
6. สิ้นสุดการทำงาน (End of Service Run)

### 3.2.3 เทคนิคการรีเจนเนอเรต

วิธีการรีเจนเนอเรตแบ่งตามทิศทางการไหลได้ 2 วิธี คือ

1. สารรีเจนเนอแรนต์ที่ไหลในทิศทางเดียวกับน้ำเสีย (Cocurrent Operation) โดยผ่านเข้าทางด้านบนและไหลออกด้านล่างของคอลัมน์ เป็นวิธีที่ง่ายและสามารถรีเจนเนอเรตได้ง่าย แต่วิธีนี้ต้องใช้สารเคมีมาก ความเข้มข้นสูง และใช้เวลานาน รวมทั้งเกิดปัญหาการรั่วไหลของไอออน ซึ่งเกิดจากไอออนที่รีเจนเนอเรตไม่หมด เคลื่อนลงสู่ด้านล่างของคอลัมน์ และหลุดออกมาเมื่อทำการผ่านน้ำเสียในการทำงานรอบต่อไป
2. สารรีเจนเนอแรนต์ที่ไหลสวนทางกับน้ำเสีย (Counter-current Operation) วิธีนี้ทำให้เกิดการรีเจนเนอเรต อย่างสมบูรณ์ มีการรั่วไหลของไอออนที่ไม่ต้องการน้อย ใช้สารเคมีอย่างมีประสิทธิภาพมากกว่าและประหยัดสารรีเจนเนอแรนต์ได้กว่าแบบไหลในทิศทางเดียวกันถึงร้อยละ 50 เหมาะสำหรับระบบที่ต้องการความบริสุทธิ์ของน้ำสูง

### 3.2.4 Purolite Resin ion exchange

เรซินที่นำมาใช้ในการทดลองเป็นของบริษัทเบนเมเยอร์จำกัด ซึ่งใช้เรซิน 2 ชนิดด้วยกันคือ  
 1. Cation C-150 และ 2. Anion A-500 ซึ่งโครงสร้างทั้งสองเป็นแบบแมคโคพอร์รัส (Macroporous Type) ซึ่งจะแตกต่างจากการทดลองแบบเดิมที่อาหาร วุฒิสถียวงศ์กุล (2545) ได้ออกแบบไว้ซึ่งใช้ เรซินในแบบเจล (Gel Type) คุณสมบัติต่างๆพิจารณาจากตารางที่ 3.1

ตารางที่ 3.1 สมบัติของเรซินที่ใช้ทดลองทั้ง 2 ชนิด\*

	Purolite C-150	Purolite A-500
Matrix	Polystyrene divinylbenzene	
Appearance	Spherical beads	Opaque, spherical beads
Functional groups	$R-SO_3^-$	$R-N(CH_3)_3^+$
Ionic form as ship	Sodium - $Na^+$	Chloride - $Cl^-$
Total exchange capacity**	1.8 eq/l minimum	1.15 eq./l minimum
Moisture Retention	48-53%	53-58%
Particle Size Range	0.3 – 1.2 mm	0.3 – 1.2 mm
Screen Size Range	16-50 US Standard Mesh	16 - 50 US Standard Mesh
Max reversible swelling	$Na^+ \Rightarrow H^+$ 5%	$Cl^- \Rightarrow OH^-$ 15%
pH Limitations	None	0 - 14

\* อ้างอิงจาก เอกสารข้อมูลเทคนิคจากบริษัทเบนเมเยอร์จำกัด

\*\* แบบเจล มีความจุของการแลกเปลี่ยนไอออน ไอออนบวก 2.0 eq/l ไอออนลบ 1.3 eq/l  
 ข้อดี และข้อเสียของเรซินแบบแมคโคพอร์รัส เปรียบเทียบกับเจล  
 ข้อดี

1. มีเสถียรภาพเชิงกลที่ดีกว่า เนื่องจากมีอันดับการเชื่อมขวางที่สูงกว่า สามารถทนการสึกกร่อนได้ดี และทนต่อปฏิกิริยาออกซิเดชัน
2. มีขนาดของรูพรุนใหญ่กว่า โดยเรซินแบบแมคโคพอร์รัส (50 -100nm) และแบบเจล (6-20nm)

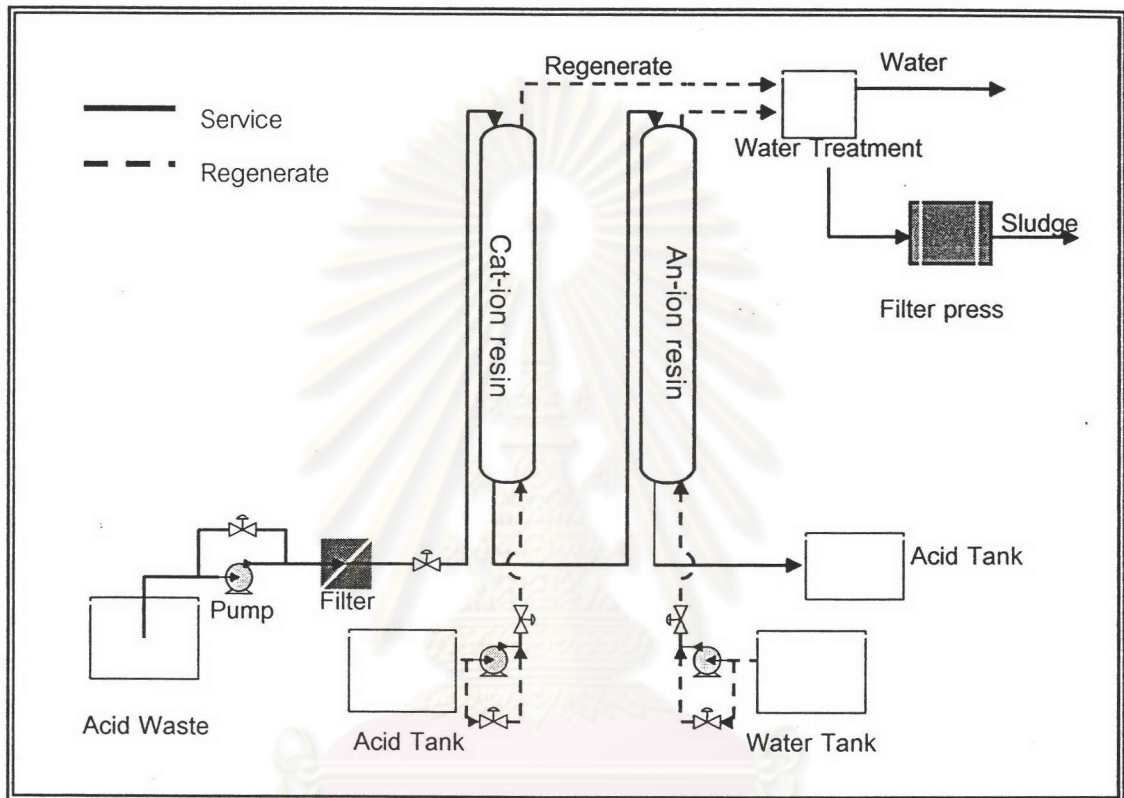
ข้อเสีย

1. การสังเคราะห์ต้องใช้ปริมาณไดไวนิลเบนซีน มากกว่า ทำให้ต้นทุนในการใช้งานสูง
2. ความจุของการแลกเปลี่ยนไอออนต่ำกว่าแบบเจล



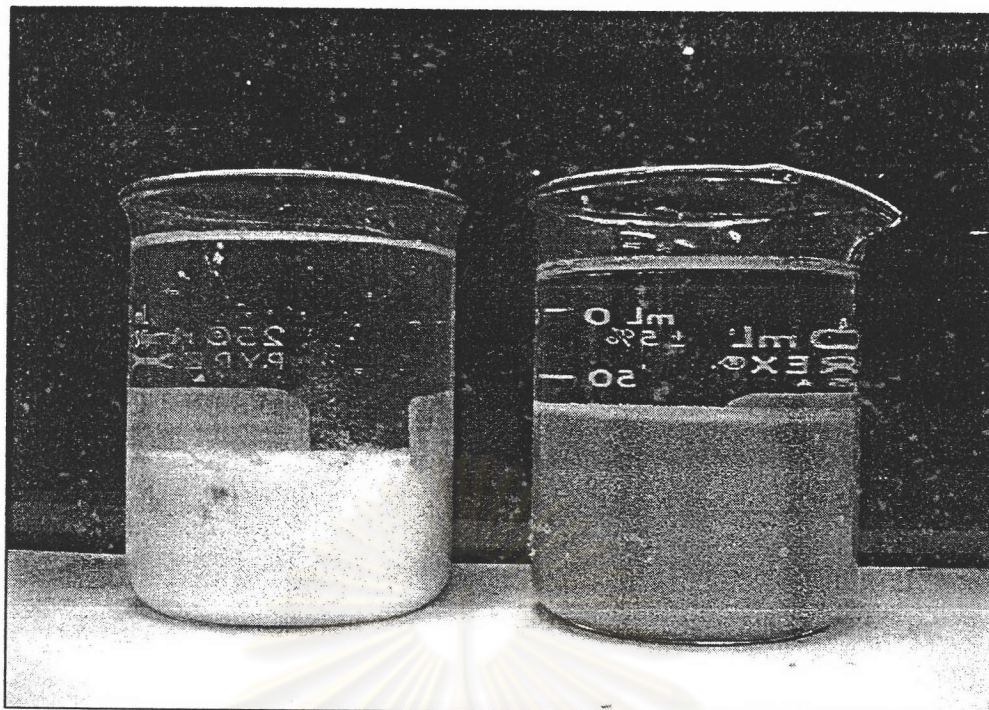
### 3.2.5 กระบวนการแลกเปลี่ยนไอออนของ อาหาร วุฒิสัตย์วงศ์กุล (2545)

มีลักษณะเป็นการนำกลับกรดด้วยวิธีแลกเปลี่ยนไอออน โดยมีการออกแบบ และสร้าง หอเรซินแบบคอลัมน์ (Fixed-Bed Column) ซึ่งประกอบด้วย แคทไอออน (Cation) และ แอนไอออน (Anion) โดยแต่ละหอบรรจุเรซินประมาณ 13 ลิตรแสดงดังรูปที่ 3.6

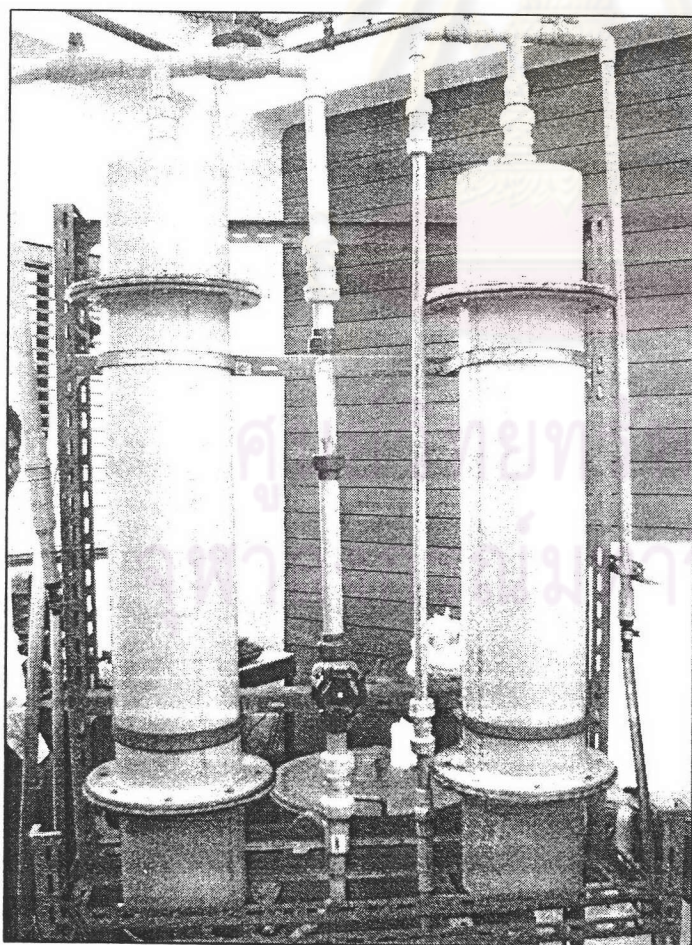


รูปที่ 3.6 แสดงกระบวนการแลกเปลี่ยนไอออนของเหล็ก (อาหาร 2545)

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.7 แสดงเรซินแลกเปลี่ยนประจุลบ (ซ้าย) และเรซินแลกเปลี่ยนประจุบวก (ขวา)



รูปที่ 3.8 แสดงหอเรซินแลกเปลี่ยนไอออน

จากรูปที่ 3.8 แสดงต้นแบบ  
กระบวนการนำกลีบริดไฮโดรคลอ  
ริกแลกเปลี่ยนไอออนมีรายละเอียด  
ดังนี้

1. หอด้านซ้ายบรรจุเรซินแลกเปลี่ยนไอออนประจุบวก Purolite C-150
2. หอด้านขวารับเรซินแลกเปลี่ยนไอออนประจุลบ Purolite A-500
3. อุปกรณ์วัดอัตราการไหล
4. อุปกรณ์ควบคุมอัตราการไหลแบบไดอะแฟรม
5. บั้มสำหรับกรดเสีย



### 3.3 กระบวนการดูดซับ (Adsorption Process)

การดูดซับเป็นกระบวนการหนึ่งที่ทำหน้าที่แยกองค์ประกอบของ ก๊าซ และของเหลว โดยมีตัวดูดซับไว้ที่ผิวของของแข็ง (Adsorbent) โดยทั่วไปแล้วของแข็งที่จะนำมาใช้เป็นตัวดูดซับต้องมีขนาดเล็ก และมีพื้นที่ผิวสำหรับดูดซับมาก โดยบรรจุเป็นชั้นตัวดูดซับในคอลัมน์ ซึ่งของไหลที่มีสารถูกดูดซับจะไหลผ่านชั้นของตัวดูดซับนี้ และเกิดการดูดซับกันระหว่าง องค์ประกอบที่ต้องการดูดซับ กับตัวดูดซับ จนกระทั่งตัวดูดซับอิ่มตัวและจะไม่สามารถดูดซับได้อีก จึงต้องมีการฟื้นฟูสภาพของตัวดูดซับ (Regenerate) เพื่อให้สามารถกลับมาใช้ได้ อีก ซึ่งกระบวนการดูดซับจะมีลักษณะเป็นรอบของการดูดซับ ตั้งแต่การเริ่มใช้งาน ตัวดูดซับอิ่มตัว และสุดท้ายทำการฟื้นฟูสภาพ และกลับมาใช้อีกครั้งหนึ่ง

ลักษณะการใช้งานของการดูดซับ เช่นการดูดซับในสถานะของเหลว ได้แก่ การแยกองค์ประกอบอินทรีย์ออกจากน้ำ หรือสารละลายอินทรีย์ การแยกสีออกจากสารอินทรีย์ และการแลกเปลี่ยนไอออนก็เป็นส่วนหนึ่งของการดูดซับ โดยการแลกเปลี่ยนไอออนเป็นลักษณะของประจุในสารละลาย หรือธาตุ แลกเปลี่ยนกับ ตัวดูดซับคือ เรซินที่มีทั้งการแลกเปลี่ยนประจุบวก และประจุลบ เป็นต้น และท้ายสุดการดูดซับในสถานะก๊าซ ได้แก่การแยกน้ำออกจาก ก๊าซ ไฮโดรคาร์บอน การแยกซัลเฟอร์ออกจากก๊าซธรรมชาติ หรือการแยกกลิ่นออกจากก๊าซ

โดยทั่วไปกระบวนการดูดซับที่นิยมใช้มี 2 ระบบได้แก่ ระบบแบท (Batch system) และระบบแพคเบด (Packed bed system) ส่วนการกำหนดอัตราการดูดซับขึ้นอยู่กับ การต้านทานการถ่ายเทมวล และ การถ่ายเทความร้อน (mass and heat transfer resistance) ในที่นี้ขอกล่าวเฉพาะกรณีที่คุณสมบัติไม่มีผลต่ออัตราการดูดซับ (Isothermal system) และระบบดูดซับหนึ่งองค์ประกอบ (Single-Component Sorption) เท่านั้น

#### 3.3.1 จลนศาสตร์ของการดูดซับในระบบแบท (Kinetics of Sorption in Batch System)

เมื่อก้าวถึงการเปลี่ยนแปลงการดูดซับในระบบต่างๆ ซึ่งมีทั้งระบบแบทซึ่งง่ายต่อการวิเคราะห์ และอธิบายกว่าระบบแพคเบดที่มีความซับซ้อนมากกว่าจะกล่าวในหัวข้อต่อไป การต้านทานการถ่ายเทมวลในระบบแบทนี้ขึ้นกับ ตัวดูดซับ ซึ่งหลักๆ แล้วประกอบด้วย ส่วนที่เป็นผิวรูพรุน (Microporous Surface Diffusion) และส่วนที่โพรงรูพรุน (Macroporous Pore Diffusion) ซึ่งการต้านทานการถ่ายเทมวลโดยแบบจำลองมีลักษณะเป็นดังนี้



$$\text{Surface Diffusion: } \frac{\partial q}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 D_c \frac{\partial q}{\partial r} \right) \quad \text{สมการ 3.3}$$

$$\text{Pore Diffusion: } (1 - \varepsilon_p) \frac{\partial q}{\partial t} + \varepsilon_p \frac{\partial c}{\partial t} = \varepsilon_p D_p \left( \frac{\partial^2 c}{\partial R^2} + \frac{2}{R} \frac{\partial c}{\partial R} \right) \quad \text{สมการ 3.4}$$

สำหรับ Linear Equilibrium

โดย  $D_c$  ค่าสัมประสิทธิ์การแพร่ของผิวรูพรุน;  $D_p$  ค่าสัมประสิทธิ์การแพร่ของโพรงรูพรุน;  $q(r,t)$  คือความเข้มข้นสารถูกดูดซับในชั้นตัวดูดซับ;  $c(R,t)$  คือความเข้มข้นสารถูกดูดซับในชั้นของเหลว;  $\varepsilon_p$  คือสัดส่วนช่องว่างของโพรงรูพรุน ซึ่งอัตราการดูดซับขึ้นอยู่กับว่าชั้นการกำหนดอัตราจะเกิดขึ้นที่ส่วนไหน ซึ่งปกติแล้วไม่สามารถสรุปได้แน่ชัดว่าเป็นแบบใด นอกจากใช้ผลการทดลองเปรียบเทียบ

### 3.3.2 จลนศาสตร์ของการดูดซับในระบบแพคเบด (Kinetics of Sorption in Packed-Bed System)

เมื่อก้าวถึงการเปลี่ยนแปลงการดูดซับในระบบนี้มีความซับซ้อนมากกว่า โดยมีขั้นตอนตั้งแต่ชั้นของไหล (Fluid Phase) ไหลเข้าสู่แพคเบดคอลัมน์ ซึ่งเราสามารถพิจารณาตัวแปรต่าง ดังนี้ ขนาดของตัวดูดซับ อัตราการไหลของ ขนาดของชั้นเบดที่ต้องนำมาหาความสัมพันธ์กับความดันในคอลัมน์ ซึ่งต่างก็มีผลกระทบต่ออาการออกแบบ ทั้งการเลือกขนาดบีบ การไหลที่ทำให้เกิดการผสมที่ดี (Well Mix) เพื่อที่จะนำมาซึ่งการดูดซับที่มีประสิทธิภาพสูงสุด ซึ่งทำให้ดูแล้วมีความซับซ้อนกว่าแบบเบท ที่พิจารณาอัตราการดูดซับจากข้อมูลของตัวดูดซับ และเวลาสัมผัสเท่านั้น ซึ่งการเปลี่ยนแปลงการดูดซับแบบแพคเบดจึงมีความซับซ้อนสูงตัวอย่างแบบจำลองทางคณิตศาสตร์ของการดูดซับในระบบแพคเบดที่นิยมนำมาใช้ในการวิเคราะห์พฤติกรรมของการดูดซับมาที่สุดเป็นดังนี้

$$\frac{\partial c}{\partial t} + \left( \frac{1 - \varepsilon}{\varepsilon} \right) \frac{\partial q}{\partial t} = - \frac{F_L}{\varepsilon A_c} \frac{\partial c}{\partial z} + D_L \frac{\partial^2 c}{\partial z^2} \quad \text{สมการ 3.5}$$

โดยที่  $\varepsilon$  คือสัดส่วนรูพรุนภายนอก (external void fraction);  $F_L$  คืออัตราการไหล;  $A_c$  คือพื้นที่หน้าตัดของคอลัมน์;  $c$  คือความเข้มข้นของสารถูกดูดซับในชั้นของไหล;  $q$  คือความเข้มข้นของสารถูกดูดซับในชั้นตัวดูดซับ;  $D_L$  คือสัมประสิทธิ์การแพร่ในชั้นของไหล ( $m^2/s$ );  $t$  คือเวลาในการดูดซับ และ  $z$  คือระยะความสูงของคอลัมน์และในพจน์แรกแสดงถึงการสะสมของสารถูกดูดซับในชั้นของของไหล พจน์ที่สองแสดงถึงการสะสมของสารถูกดูดซับในชั้นของตัวดูดซับ พจน์ที่สามแสดงถึงปริมาณของสารถูกดูดซับที่เคลื่อนที่และไหลผ่านชั้นตัวดูดซับ พจน์สุดท้ายแสดงถึงการแพร่ของสารถูกดูดซับในชั้นตัวดูดซับ

ในส่วนพจน์ที่สองสามารถพิจารณาสมการของความสมดุลระหว่างสารถูกดูดซับในชั้นของของไหล  $q$  และสารถูกดูดซับในชั้นของตัวดูดซับ  $c$  ได้จาก การต้านทานการถ่ายเทมวลของตัวดูดซับ (Mass Transfer Resistance of Adsorbent Particles) ดังสมการนี้

$$\frac{\partial q}{\partial t} = \frac{3}{R_p} k_f (c - c_p) \quad \text{สมการ 3.6}$$

โดยที่  $k_f$  คือสัมประสิทธิ์อัตราการถ่ายเทมวล;  $c_p$  คือความเข้มข้นของสารถูกดูดซับบริเวณผิวรูพรุนของเรซิน;  $R_p$  คือรัศมีของเม็ดเรซิน นำไปแทนลงในสมการ 3.5 จะได้

$$\frac{\partial c}{\partial t} + \left( \frac{1-\varepsilon}{\varepsilon} \right) \frac{3}{R_p} k_f (c - c_p) = -\frac{F_L}{\varepsilon A_c} \frac{\partial c}{\partial z} + D_L \frac{\partial^2 c}{\partial z^2} \quad \text{สมการ 3.7}$$

จากสมการนี้เราสามารถหาโพลีความเข้มข้น (Concentration profile) และสามารถนำไปแก้ปัญหาสมการอนุพันธ์ย่อย ซึ่งสามารถกำหนดเงื่อนไขขอบได้ สามารถแก้ปัญหาโดยวิธีการทางตัวเลข แต่การจำลองด้วยวิธีทางตัวเลขนี้จำเป็นต้องรู้ข้อมูลของค่าสัมประสิทธิ์ต่างๆ ในแบบจำลองเช่น  $\varepsilon; k_f; D_L$  เป็นต้น รายละเอียดการทำสมดุลมวลสารของการดูดซับ (ภาคผนวก ก)

### 3.3.3 สมบัติทางกายภาพของตัวดูดซับ (Adsorbent)

โดยทั่วไปตัวดูดซับ (Adsorbent) มีการพัฒนาเพื่อให้สามารถใช้งานกับกระบวนการแยก (Separations) ได้หลากหลาย ตัวดูดซับส่วนใหญ่มีลักษณะเป็นลูกกลมรูพรุนขนาดเล็ก (pellets, beads, granules) ตัวดูดซับนั้นมีโครงสร้างที่มีปริมาณรูพรุนสูง และสัดส่วนปริมาณรูพรุนต่อปริมาณตัวดูดซับอย่างน้อย 50% ซึ่งประเภทของการดูดซับ แบ่งได้ดังนี้ การดูดซับทางกายภาพ (Physical Adsorption) เป็นการดูดซับโดยอาศัยแรงยึดเหนี่ยวระหว่างโมเลกุล ได้แก่แรงแวนเดอร์วาล (Van der Waals force) และการดูดซับทางเคมี (Chemical Adsorption) เป็นการดูดซับโดยอาศัยแรงยึดเหนี่ยวที่เกิดจากการสร้างพันธะทางเคมี ได้แก่แรงดึงดูดทางเคมี (Chemical interaction) และการดูดซับแบบการแลกเปลี่ยนไอออนที่เกี่ยวข้องกับงานวิจัย โดยปฏิกิริยาทั้งสองเป็นแบบคายความร้อน (Exothermic reaction) ส่วนลำดับขั้นตอนในการดูดซับเริ่มจาก

- สารถูกดูดซับ (Solute) แพร่จากชั้นของไหล (Bulk fluid) เข้าไปใน ผิวนอกของตัวดูดซับ
- สารถูกดูดซับ (Solute) แพร่จากผิวนอกของตัวดูดซับเข้าสู่โพรงรูพรุน
- สุดท้ายสารถูกดูดซับ (Solute) ถูกดูดซับไว้ที่ผิวภายในของรูพรุน

### 3.3.4 ภาวะสมมูลของการดูดซับ (Adsorption Equilibrium)

ภาวะสมมูลระหว่างชั้นของเหลวและชั้นของตัวดูดซับ สำหรับการดูดซับองค์ประกอบเดียว (Single component adsorption) หลายองค์ประกอบ (Multicomponent adsorption) หรือการแลกเปลี่ยนไอออน (Ion Exchange) นั้นเป็นปัจจัยหนึ่งที่มีความสำคัญในสำหรับบ่งบอกถึงประสิทธิภาพการดูดซับ และในสภาวะที่อุณหภูมิของกระบวนการคงที่ ซึ่งเรียกว่า การดูดซับแบบอุณหภูมิคงที่ (Adsorption Isotherm) ซึ่งมีความสัมพันธ์ดังนี้

$$q = f(c, T) \quad \text{สำหรับการดูดซับสภาวะของเหลว}$$

$$q = f(c) \quad \text{สำหรับการดูดซับสภาวะของเหลวแบบอุณหภูมิคงที่}$$

โดยที่	$q$	คือความเข้มข้นของสารถูกดูดซับในชั้นตัวดูดซับ ปริมาณสารถูกดูดซับ / ปริมาณตัวดูดซับ
	$c$	คือความเข้มข้นของสารถูกดูดซับในสภาวะของเหลว ปริมาณสารถูกดูดซับ / ปริมาณของเหลว
	$T$	คืออุณหภูมิ

สามารถแบ่งภาวะสมมูลแบบอุณหภูมิคงที่ได้ดังนี้

- Linear Equilibrium Isotherm  $q = Kc$  สมการ 3.8

ซึ่งเป็นการประมาณอย่างง่ายที่สุดของภาวะสมมูลระหว่างเฟสของเหลว และเฟสของตัวดูดซับมีลักษณะเป็นสมมูลแบบเชิงเส้นดังนี้ โดย  $K$  คือค่าคงที่ที่สามารถหาได้จากการทดลอง ซึ่งไม่สามารถจะนำไปใช้งานจริงได้จึงมีการเพิ่มความซับซ้อนของภาวะสมมูลนี้

- Freundlich Equilibrium Isotherm  $q = Kc^n$  สมการ 3.9

เมื่อพิจารณาจากสมการแล้วจะมีความซับซ้อนเพิ่มมากขึ้น โดย  $K$ ,  $n$  คือค่าคงที่ที่สามารถหาได้จากการทดลอง

- Langmuir Equilibrium Isotherm  $q = \frac{q_0 c}{K + c}$  สมการ 3.10

โดย  $q_0$ ,  $K$  คือค่าคงที่ที่สามารถหาได้จากการทดลอง



### 3.3.5 การดูดซับไอออนในสารละลายเจือน้ำ (Adsorption of ion from aqueous solution)

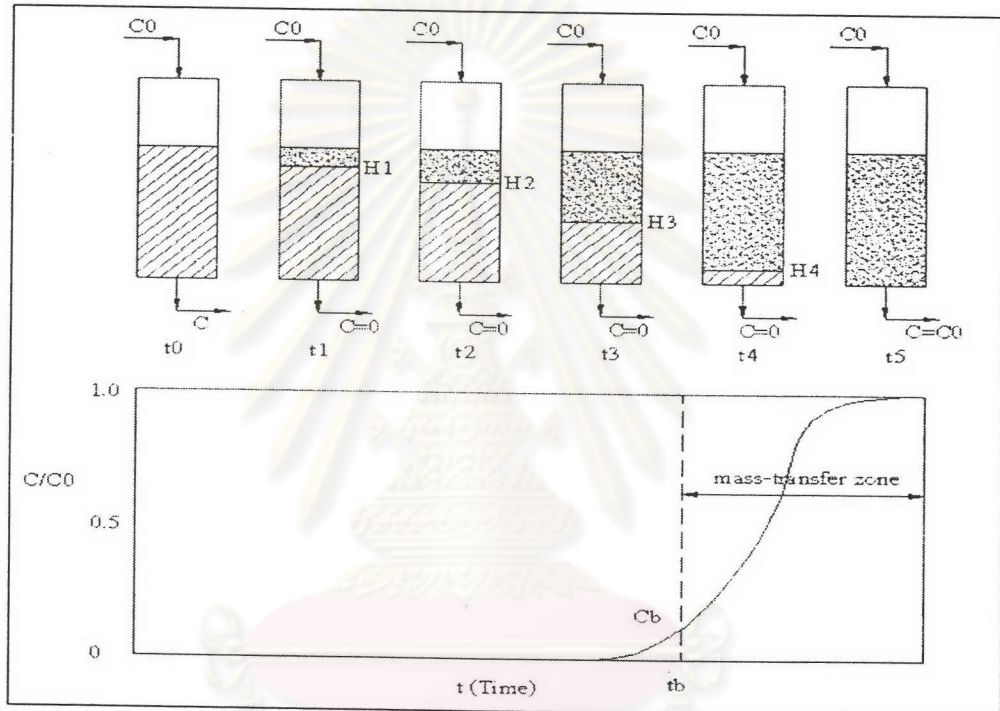
ไอออนของโลหะซึ่งมีทั้งไอออนบวก (Cation) และ ไอออนลบ (Anion) โดยเฉพาะไอออนลบมักจะรวมตัวเป็นแบบเชิงซ้อน (Complex forms) ในสายละลาย ซึ่งส่วนใหญ่มีลักษณะของภาวะสมดุลแบบ Freundlich และ Langmuir isotherm (Rubin และ Mercer 1981) เนื่องด้วยธรรมชาติแล้วการแลกเปลี่ยนไอออนมีลักษณะการดูดซับเชิงเคมี มากกว่า การดูดซับทางกายภาพ ดังนั้นการพิจารณาถึงแบบจำลองของการแลกเปลี่ยนไอออนนี้จึงควรมีลักษณะเป็นแบบซับซ้อนสูง (Complexation Model) ดังนั้นการระบุแบบจำลองของการดูดซับด้วยการแลกเปลี่ยนไอออนจึงมีความยากด้วยเราไม่สามารถที่จะทราบถึงธรรมชาติการดูดซับที่เกิดขึ้นภายในพื้นผิวโพรงรูพรุนอย่างไรก็ตามก็ยังมีแนวคิดแบบจำลองแบบต่างๆ มากมายออกมาให้ทำการศึกษาพฤติกรรม การดูดซับโดยทั่วไปใช้แบบจำลองทางคณิตศาสตร์แบบพื้นฐานมาวิเคราะห์เพื่อหาค่าตัวแปร เช่น ความเข้มข้นไอออนชนิดต่างๆ หรือแม้แต่โฟโพลีความเข้มข้นเป็นต้น

### 3.3.6 พลศาสตร์การดูดซับของฟีกเบดคอลัมน์ และ โฟโพลีความเข้มข้น

(Dynamic of Adsorption Fixed bed Columns and Concentration Profile)

จากการเปลี่ยนแปลงพฤติกรรมของฟีกเบดคอลัมน์นั้นเกี่ยวข้องกับความสัมพันธ์ของตัวดูดซับ คือในกระบวนการต้องมีการดูดซับ (Adsorption) และการฟื้นฟูสภาพ (Desorption) โดยความเข้มข้นขาออก (Effluent Concentration) จากคอลัมน์นั้นพิจารณาจาก ผลของการกระจายตามแนวแกน (Axial Dispersion) และ การต้านทานการถ่ายเทมวล (Mass transfer resistance) โดยผลของความเข้มข้นขาออกที่ได้จากการทดลอง (Experimental Response) ต้องใกล้เคียงกับ ผลของความเข้มข้นจากทฤษฎี (Theoretical Response) ซึ่งในส่วนของทฤษฎีนั้นได้มาจากแบบจำลองทางคณิตศาสตร์ของกระบวนการ ในทางกลับกัน การทำนายโฟโพลีความเข้มข้น จากข้อมูลจลนศาสตร์พื้นฐานและภาวะสมมูลก็เป็นวิธีหนึ่งที่ใช้ในการหาประสิทธิภาพการดูดซับของคอลัมน์นั้นๆ ได้ ส่วนการพิจารณาพฤติกรรมการดูดซับ และกราฟระหว่างความเข้มข้น กับเวลา หรือ โฟโพลีความเข้มข้น  $\frac{C}{C_0}$  โดย  $C$  คือความเข้มข้นสารถูกดูดซับขาออก และ  $C_0$  คือความเข้มข้นสารถูกดูดซับขาเข้า ที่เรียกว่า Breakthrough Curve ซึ่งสามารถนำมาใช้ในการออกแบบระบบดูดซับ และใช้ในการหาสถานะที่เหมาะสมในการดำเนินการ ดังนั้นการทำนายโฟโพลีความเข้มข้นจึงมีความสำคัญมาก ในการบ่งบอกถึงประสิทธิภาพการดูดซับ และช่วยหาสถานะในการดำเนินการ หรือการลดพลังงานในการฟื้นฟูสภาพตัวดูดซับได้

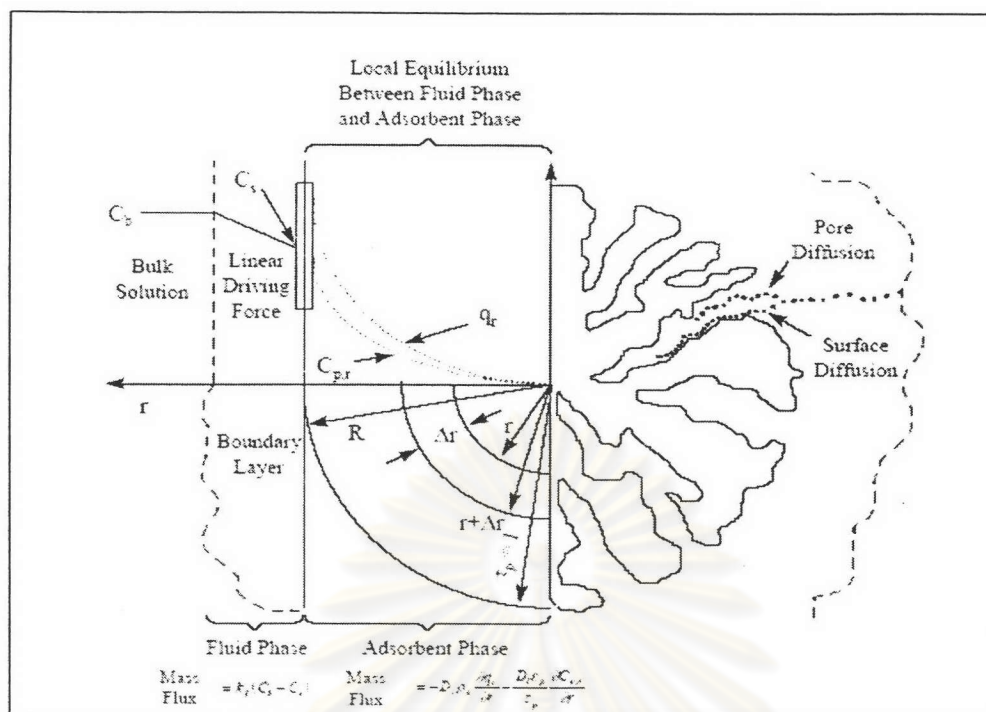
เมื่อก้าวถึงโพลีความเข้มข้นนั้น สามารถพิจารณาจากการทดลองโดยเริ่มจาก ให้ ความเข้มข้นสารถูกดูดซับในชั้นเบดขณะเริ่มต้นเป็นศูนย์  $c = 0$  ดังนั้น  $\frac{c}{c_0} = 0$  เมื่อพิจารณาเมื่อ เวลาผ่านไป  $t_1$  ในชั้นเบด  $H_1$  ก็อิ่มตัว แต่ความเข้มข้นขาออกยังคงมีค่าเป็นศูนย์  $c = 0$  เพราะ ในชั้นเบดยังสามารถถ่ายเท และดูดซับได้ และเมื่อเวลาผ่านไป  $t_1 - t_2$  ชั้นเบดที่อิ่มตัวเพิ่มขึ้นตั้งแต่  $H_1 - H_2$  ซึ่งจากภาพชั้นเบดเริ่มอิ่มตัวด้วยสารถูกดูดซับแล้ว ทำให้ ณ เวลา  $t_3$  ความเข้มข้นสาร ถูกดูดซับค่อยๆ ออกมาจากคอลัมน์ที่เพิ่มขึ้นอย่างรวดเร็วซึ่งทำให้เกิดเป็น Breakthrough Curve ดังรูปที่ 3.9



รูปที่ 3.9 แสดงลักษณะกราฟ Breakthrough Curve

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย





รูปที่ 3.10 แสดงกลไกการดูดซับ

### 3.3.7 กลไกการถ่ายเท และการแพร่ (Transport and Dispersion Mechanisms)

ในส่วนนี้จะพิจารณาการถ่ายเทและการแพร่ในเม็ดเรซิน ซึ่งแต่ละขั้นตอนของการดูดซับที่ เคยกล่าวไว้แล้วนำทฤษฎีส่วนนี้ในการอธิบาย ซึ่งสามารถแบ่งได้เป็น 2 ส่วนดังนี้

#### 3.3.7.1 กลไกการถ่ายเทภายในตัวดูดซับ (Intraparticle Transport Mechanisms) ประกอบด้วย

1. การแพร่เข้าสู่โพรงรูพรุน (Pore diffusion) เป็นลักษณะของสารถูกดูดซับกำลังเคลื่อนที่ และแพร่เข้าสู่โพรงรูพรุนของตัวดูดซับ ส่วนแรงผลักดันเคลื่อน (driving force) ในการแพร่ผ่านนี้ สามารถประมาณได้จากเกรเดียนต์ความเข้มข้นของสารถูกดูดซับในโพรงรูพรุน

2. การแพร่เข้าสู่ผิวรูพรุน (Surface diffusion) เป็นลักษณะของสารถูกดูดซับกำลังเคลื่อนที่เข้าไปแลกเปลี่ยน หรือดูดซับเอาไว้บริเวณผิวรูพรุน ส่วนแรงผลักดันเคลื่อน (driving force) ในการแพร่ผ่านนี้สามารถประมาณได้จากเกรเดียนต์ความเข้มข้นของสารถูกดูดซับในผิวรูพรุน

3. ปฏิกิริยาบริเวณผิวตัวดูดซับ (Reaction kinetics at Phase Boundaries) เป็นลักษณะ การเกิดปฏิกิริยาของการดูดซับบริเวณผิว ซึ่งจะเกิดขึ้นเฉพาะกรณี เช่น ในกระบวนการแยกทางชีวภาพ (Biological separation)

### 3.3.7.2 กลไกการถ่ายเท และการแพร่ภายนอกตัวดูดซับ

(Extraparticle Transport and Dispersion Mechanisms)

4. การถ่ายเทมวลภายนอก (External mass transfer) หรือเรียกว่า Film diffusion เป็นลักษณะของการถ่ายเทข้ามชั้นระหว่าง สารถูกดูดซับ กับชั้นของไหล (Bulk fluid)

5. กลไกการผสม (Mixing) เป็นลักษณะของการสารถูกดูดซับกระจายอยู่เต็มภายในคอลัมน์เฮลด์การเกิดช่องว่างการดูดซับ (Dead zones) ดังนั้นรูปแบบการไหล และอัตราการไหลที่เหมาะสมจึงสามารถเพิ่มประสิทธิภาพของดูดซับได้

## 3.4 ทฤษฎีข่ายงานนิวรัล

### 3.4.1 วิวัฒนาการของเซลล์ประสาท

การศึกษาเกี่ยวกับโครงสร้างระบบประสาทของมนุษย์ (Human Neurons System) มีมาเป็นเวลานาน นั่นคือระบบประสาทของมนุษย์ประกอบด้วยกลุ่มของเซลล์ประสาทประมาณ  $10^{11}$  เซลล์ โดยสามารถติดต่อกับเซลล์ประสาทอื่นๆ โดยผ่าน แอกซอน (Axon) และซินแนปส์ (Synapses) ซึ่งความหนาแน่นของซินแนปส์ประมาณ  $10^4$  หน่วยต่อเซลล์ประสาท สมมติฐานเกี่ยวกับแบบจำลองของระบบประสาทคือเซลล์ประสาทจะทำการติดต่อกันโดยอาศัยการกระตุ้นทางไฟฟ้า (Electrical Impulse) แต่การทำงานของเซลล์ประสาทนั้นเกิดจากกระบวนการทางชีวเคมี ซึ่งช่วงเวลาในการประมวลผลของเซลล์ประสาท (Neuron Switch Time) จะมีค่าน้อยกว่าหนึ่งในพันวินาทีซึ่งน้อยกว่าเวลาที่กระแสไฟฟ้าวิ่งผ่านคอมพิวเตอร์ถึงหนึ่งล้านเท่าและเซลล์ประสาทสามารถเชื่อมโยงกันได้ไวกว่าซูเปอร์คอมพิวเตอร์ในปัจจุบันถึงหนึ่งพันเท่า

ในทางชีววิทยาของเซลล์ประสาทมีองค์ประกอบ 3 ส่วนดังนี้

1. เดนไดรท์ (Dendrite) ที่แตกแขนงออกจากตัวเซลล์เพื่อรับสัญญาณจากแอกซอนของเซลล์ประสาทตัวอื่นเข้าและส่งต่อมายังเซลล์ประสาทถัดไป
2. ตัวเซลล์ประสาท (Cell Body) ทำหน้าที่รวบรวมสัญญาณจากเดนไดรท์ และส่งไปให้แอกซอน
3. แอกซอน (Axon) ทำหน้าที่นำสัญญาณที่ออกจากเซลล์ประสาทส่งไปยังเดนไดรท์ของเซลล์ประสาทอื่นๆ

ด้วยความสนใจของนักวิทยาศาสตร์ที่ต้องการเลียนแบบกระบวนการคิดของสมอง และเซลล์ประสาท จึงได้พยายามศึกษารูปแบบ และกระบวนการคิดเหล่านั้นซึ่งเป็นเรื่องที่ยากมากในการทำ ความเข้าใจกระบวนการคิดของสมอง แต่เป็นก้าวแรกของการศึกษาข่ายงานประสาทเทียม (Artificial Neural Network) เริ่มจากงานวิจัยของ Warren McCulloch นักประสาทวิทยา และ

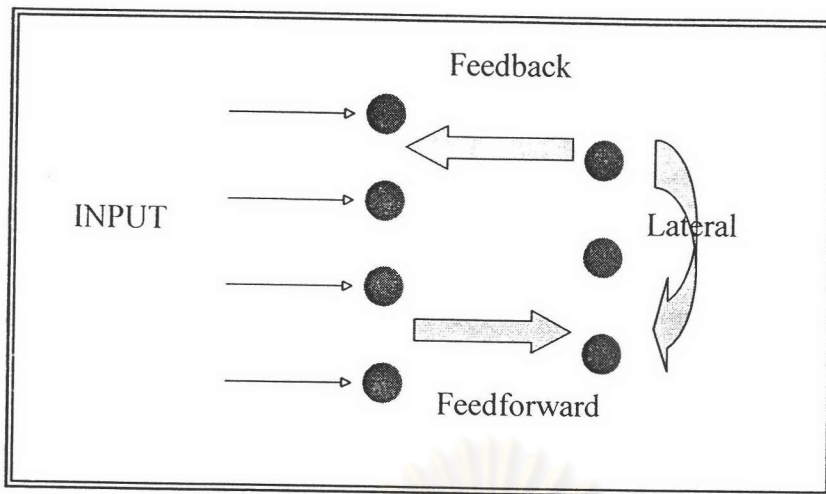


Walter Pitts นักคณิตศาสตร์ (1943) มีการนำเสนอข่ายงานที่จำลองลักษณะการเรียนรู้ของเซลล์ประสาทที่เรียกว่า “Hebbian learning” โดยใช้สมการคณิตศาสตร์อย่างง่าย ต่อจากนั้นในยุค 1950s จึงได้มีการนำเครื่องคอมพิวเตอร์มาใช้ในการคำนวณ โดย Nathaniel Rochester (1956) จากศูนย์วิจัยของ IBM ซึ่งเป็นการเริ่มต้นอย่างจริงจัง ต่อมาได้มีการคิดค้นข่ายงานชนิดต่างๆ และพัฒนาจนกระทั่งในปี 1958 Frank Rosenblatt นักประสาท-ชีววิทยาได้มีการเสนอข่ายงาน “เปอร์เซปตรอน (Perceptron)” ซึ่งมีพื้นฐานการทำงานเป็น 3 ชั้น คือ ชั้นอินพุต ชั้นเอาต์พุต และชั้นซ่อน ต่อมา Bernard Widrow และ Marcian Hoff (1959) ได้พัฒนาข่ายงาน ADALINE (ADaptive Linear Neuron) และได้รับการปรับปรุงเป็นข่ายงาน MADALINE (Multiple ADALINE) เหล่านี้เป็นข่ายงานที่สามารถแก้ปัญหาได้เฉพาะแบบเชิงเส้น เนื่องจากมีชั้นของซ่อนเพียงชั้นเดียว เพื่อแก้ปัญหาแบบไม่เป็นเชิงเส้นจึงมีการคิดค้นข่ายงานนิเวรลแบบการกระจายย้อนกลับ (Backpropagation Neural Network) โดย Rumelhart (1986) ซึ่งเป็นวิธีการปรับค่าถ่วงน้ำหนักจากชั้นเอาต์พุต ไปยังชั้นอินพุต ในปัจจุบันจึงมีการนำความรู้ของข่ายงานนิเวรลสามารถนำมาประยุกต์ใช้ทั้งทางด้านวิศวกรรม วิทยาศาสตร์ การแพทย์ เศรษฐศาสตร์ และการเงิน โดยงานวิจัยทั้งการส่งผ่านข้อมูล และรูปแบบการเรียนรู้

### 3.4.2 ข่ายงานนิเวรล

#### 3.4.2.1 โครงสร้างเชื่อมโยงของข่ายงานนิเวรล

1. การเชื่อมโยงแบบป้อนไปข้างหน้า (Feedforward Connection) เป็นการนำข้อมูลของนิเวรลจากชั้นที่ต่ำกว่า นำไปใช้ในชั้นที่สูงกว่า
2. การเชื่อมโยงแบบป้อนกลับ (Feedback Connection) เป็นการนำข้อมูลของนิเวรลจากชั้นที่สูงกว่า และส่งกลับนำไปใช้ในชั้นที่ต่ำกว่า
3. การเชื่อมโยงภายในชั้น (Lateral Connection) เป็นการนำข้อมูลของนิเวรลจากภายในชั้นเดียวกัน กระจายส่งกลับมาใช้ชั้นเดิม
4. การเชื่อมโยงแบบหน่วงเวลา (Time-delayed Connection) เป็นการนำข้อมูลมาหน่วงเวลาแล้วนำเข้าไปรวมกับการเชื่อมโยงของข้อมูล



● หน่วยนิวรัล

รูปที่ 3.11 แสดงโครงสร้างเชื่อมโยงของข่ายงานนิวรัล

#### 3.4.2.2 องค์ประกอบที่มีความสำคัญสำหรับการทำงานของข่ายงานนิวรัล

ในส่วนนี้กล่าวถึงองค์ประกอบทั้ง 7 ที่ใช้ออกแบบข่ายงานนิวรัลสำหรับ ชั้นอินพุต ชั้นเอาต์พุต และ ชั้นซ่อน

**องค์ประกอบที่ 1 ปัจจัยน้ำหนัก (Weight Factors):** ในข่ายงานนิวรัลจะประกอบด้วยสัญญาณอินพุตหนึ่งๆ ที่มีความสัมพันธ์กับค่าน้ำหนักหนึ่ง ซึ่งจะส่งผลไปหน่วยประมวลผล (Processing Element) คือ ฟังก์ชันมูลฐาน ซึ่งน้ำหนักนี้เป็นสัญญาณเชื่อมโยงระหว่างนิวรัลมีการปรับเปลี่ยนค่าอยู่ตลอดเวลา และในระหว่างการเรียนรู้ข่ายงานมีการสร้างแบบจำลองภายในขึ้นมาเพื่อให้มีค่าใกล้เคียงกับระบบที่สนใจ แบบจำลองภายในที่สร้างขึ้นมาจะสามารถใช้แทนแบบจำลองของกระบวนการได้ก็ต่อเมื่อผลลัพธ์ที่ออกจากระบบมีค่าใกล้เคียงกับผลลัพธ์ที่ออกจากข่ายงาน นั่นคือ ความแตกต่างของผลลัพธ์ที่ออกจากระบบกับผลลัพธ์ที่ออกจากข่ายงานต้อง มีค่าน้อยที่สุดอยู่ในช่วงที่ยอมรับได้

**องค์ประกอบที่ 2 ฟังก์ชันมูลฐาน (Basis Function):** คือการแทนการแมพฟังก์ชัน (mapping) ทางคณิตศาสตร์ด้วยฟังก์ชัน  $u(w, x)$  มีค่าเท่ากับ net เมื่อ  $w$  คือ เมทริกของน้ำหนัก และ  $x$  คือเวกเตอร์อินพุต กล่าวคือ ฟังก์ชันมูลฐานนี้เป็นการรวมสัญญาณอินพุต และน้ำหนัก โดยผ่านฟังก์ชันทางคณิตศาสตร์ ฟังก์ชันมูลฐานสามารถแบ่งเป็น 2 ประเภท คือ ฟังก์ชันมูลฐานเชิงเส้น และฟังก์ชันมูลฐานรัศมี

- ฟังก์ชันมูลฐานเชิงเส้น (Linear Basis Function)

ฟังก์ชันมูลฐานเชิงเส้นเป็นฟังก์ชันชนิดไฮเปอร์เพลน (hyperplane) โดยเป็นฟังก์ชันมูลฐานอันดับที่ 1 ค่า net ที่ได้เป็นผลรวมเชิงเส้นของค่าอินพุตซึ่งแสดงในสมการ

$$u_i(w, x) = \sum_{j=1}^n w_{ij} x_j$$

สมการ 3.11

▪ ฟังก์ชันมูลฐานรัศมี (Radial Basis Function)

ฟังก์ชันมูลฐานรัศมีเป็นฟังก์ชันชนิดไฮเปอร์สเฟียร์ (hypersphere) โดยเป็นฟังก์ชันมูลฐานอันดับที่ 2 (ไม่เชิงเส้น) ค่า net ที่ได้จะแทนระยะทางไปยังชุดข้อมูลอ้างอิง (reference pattern)

$$u_i(w, x) = \sqrt{\sum_{j=1}^n (x_j - w_{ij})^2} \quad \text{สมการ 3.12}$$

**องค์ประกอบที่ 3 ฟังก์ชันกระตุ้น (Activation Function):** ผลที่ได้จากฟังก์ชันมูลฐาน ถูกแปลงสัญญาณเพื่อให้ได้ค่าเอาต์พุต โดยฟังก์ชันกระตุ้น ซึ่งสามารถแบ่งออกเป็น ฟังก์ชันกระตุ้นแบบเชิงเส้น และฟังก์ชันกระตุ้นแบบไม่เป็นเชิงเส้น แสดงตามตารางที่ 3.2

ตารางที่ 3.2 แสดงลักษณะสมการคณิตศาสตร์ของฟังก์ชันกระตุ้นแต่ละชนิด

ชนิดฟังก์ชัน	สมการคณิตศาสตร์	ลักษณะฟังก์ชัน
1. ฟังก์ชันสเต็ป (Step Function)	$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$	แบบเชิงเส้น
2 ฟังก์ชันแรมพ์ (Ramp Function)	$f(x) = \begin{cases} 1 & \text{if } x \geq 1 \\ x & \text{if } x < 1 \\ -1 & \text{if } x < -1 \end{cases}$	แบบเชิงเส้น
3. ฟังก์ชันซิกมอยด์ (Sigmoid Function)	$f(x) = \frac{1}{1 + e^{-\frac{x}{\lambda}}}$	แบบไม่เป็นเชิงเส้น
4. ฟังก์ชันเกาส์เซียน (Gaussian Function)	$f(x) = ce^{-\frac{x^2}{\lambda^2}}$	แบบไม่เป็นเชิงเส้น
5. ฟังก์ชันอาร์คแทนเจนต์ (Arc Tangent Function)	$f(x) = \arctan\left(\frac{x}{3.1416}\right) + 0.5$	แบบไม่เป็นเชิงเส้น

เหตุผลที่ต้องใช้ฟังก์ชันกระตุ้น คือต้องการทำให้ข่ายงานเป็นแบบไม่เป็นเชิงเส้น ซึ่งถ้าเป็นแบบจำลองเชิงเส้นก็สามารถใช้ข่ายงานแบบเปอร์เซปตรอน (perceptron ที่ไม่มีชั้นซ่อน) และฟังก์ชันกระตุ้นแบบเชิงเส้นได้ แต่ถ้าต้องการแบบไม่เป็นเชิงเส้นต้องมีฟังก์ชันกระตุ้นที่แทนความไม่เป็นเชิงเส้น นั่นคือการทำข่ายงานแบบหลายชั้น และมีฟังก์ชันกระตุ้นที่ไม่เป็นเชิงเส้นเช่น ซิกมอยด์ได้แก่ logistic, tanh และ เกาส์เซียน



**องค์ประกอบที่ 4** การกำหนดขอบเขตสเกล (Scaling and Limiting): เป็นขั้นตอนการกำหนดช่วงข้อมูลด้วยการทำ Normalized เพื่อให้ข้อมูลเป็นหน่วยเดียวกันดังนี้

$$x_{nor} = \frac{x - \bar{x}}{\sigma_x}, \quad p_{nor} = \frac{p - \bar{p}}{\sigma_p} \quad y = y_{nor}\sigma_y + \bar{y}$$

โดย  $\bar{x}, \bar{p}$  และ  $\sigma_x, \sigma_p$  คือค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของข้อมูลอินพุต และเอาต์พุตของการทดลอง ส่วน  $\bar{y}, \sigma_y$  คือค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของข้อมูลเอาต์พุตจากข่ายงาน ส่วน  $x_{nor}, p_{nor}, y_{nor}$  คือค่าข้อมูลอินพุต ข้อมูลเอาต์พุตของการทดลอง และข้อมูลเอาต์พุตของข่ายงานที่ผ่านการ Normalized

**องค์ประกอบที่ 5** สัญญาณเอาต์พุต (Output Signal): จากการประมวลผล (Processing Elements) ที่ผ่านมาจะได้สัญญาณเอาต์พุต ซึ่งสมมูลกับผลของฟังก์ชันกระตุ้น

**องค์ประกอบที่ 6** ฟังก์ชันความผิดพลาด (Error Function): ในการฝึกข่ายงานค่าความแตกต่างระหว่าง เอาต์พุตปัจจุบัน และ เอาต์พุตเป้าหมาย ค่าความแตกต่างนี้จะถูกแปลงให้เป็น "ฟังก์ชันความผิดพลาด" ซึ่งมีอิทธิพลต่อโครงสร้างของข่ายงานนิวรัล ค่าความผิดพลาดเหล่านี้จะถูกกระจายไปยัง ฟังก์ชันการเรียนรู้ (Learning Function) เพื่อนำไปประมวลผลต่อไป กล่าวคือจะย้อนกลับไปปรับค่าน้ำหนัก ที่จะไปรวมกับอินพุต และจะคำนวณต่อไปเรื่อยๆ ในรอบของการเรียนรู้ เพื่อให้ได้ค่าความผิดพลาดรวมของข่ายงานที่น้อยที่สุด ซึ่งมีหลายประเภทดังนี้

- ความผิดพลาดรวมกำลังสอง (Sum Square Error)

โดยใช้ทั้งชุดข้อมูลมารวมกันจะได้เป็นผลรวมความผิดพลาดยกกำลังสอง หรือ ความผิดพลาดรวม (Sum Square Error, SSE)

$$SSE = \sum_{i=1}^N (y_i - p_i)^2 \quad \text{สมการ 3.13}$$

ซึ่ง SSE คือความผิดพลาดรวม,  $y$  คือค่าเอาต์พุตข่ายงาน และ  $p$  คือค่าเอาต์พุตการทดลอง

- ความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error)

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - p_i)^2 \quad \text{สมการ 3.14}$$

- ความผิดพลาดเฉลี่ยสมบูรณ์ (Mean Absolute Error)

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - p_i| \quad \text{สมการ 3.15}$$

**องค์ประกอบที่ 7 ฟังก์ชันการเรียนรู้ (Learning Function):** เป้าหมายหลักของฟังก์ชันการเรียนรู้ เพื่อปรับค่าน้ำหนัก ที่เป็นตัวเชื่อมโยงของอินพุต สำหรับการประมวลผลในรอบต่อไป เพื่อให้ได้ผลตามค่าเป้าหมาย สำหรับการเรียนรู้มีด้วยกัน 2 แบบ

- **ข่ายงานนิเวศแบบมีการชี้แนะ (Supervised Neural Network)**

เป็นการฝึกข่ายงานนิเวศให้เลียนแบบชุดข้อมูลตามที่กำหนดและสามารถส่งสัญญาณเอาต์พุตได้คล้ายคลึงกับชุดข้อมูลที่ป้อนให้เรียนรู้

ข่ายงานนิเวศที่มีการชี้แนะถูกฝึกโดยใช้คู่ลำดับของค่าอินพุต และค่าเป้าหมาย ค่าอินพุตใช้เป็นอินพุตให้แก่ข่ายงาน และค่าเป้าหมายใช้เป็นผู้สอน ซึ่งถือว่าเป็นข้อได้เปรียบของการเรียนรู้วิธีนี้ เพราะมีเป้าหมายที่ต้องการเป็นผู้สอนโดยค่าที่แตกต่างกันระหว่างค่าเอาต์พุตที่ออกจากข่ายงานกับค่าเป้าหมายที่ต้องการจะเป็นข้อมูลที่ใช้ในการเรียนรู้เพื่อทำให้ค่าแตกต่างระหว่างค่าเป้าหมายและเอาต์พุตของข่ายงานมีค่าน้อยที่สุด วิธีการปรับค่าน้ำหนักของข่ายงานเพื่อให้ข่ายงานนิเวศเรียนรู้และจดจำข้อมูลสัญญาณอินพุตและสร้างสัญญาณเอาต์พุตได้สอดคล้องกับข้อมูลสัญญาณเป้าหมายตามที่กำหนดในชุดของข้อมูลที่ฝึก ซึ่งกระบวนการนี้อาจทำซ้ำจนค่าผิดพลาดอยู่ในช่วงที่ยอมรับได้ ในขณะที่การเรียนรู้แบบไม่มีการชี้แนะ (Unsupervised Neural Network) จะมีเพียงชุดข้อมูลอินพุตของการฝึกเท่านั้น โดยการเรียนรู้จะเป็นการปรับตัวบนพื้นฐานของประสบการณ์เพื่อจัดกลุ่ม (Cluster) ให้เข้ากับชุดข้อมูลที่เรียนรู้ที่ผ่านมา ในทางปฏิบัตินั้นการเรียนรู้แบบชี้แนะจะสอดคล้องกับวิธีการออปติไมเซชันแบบสโตกาสติกเกรเดียนต์เดสเซนท์ (Stochastic Gradient Optimization) ตัวอย่างของข่ายงานประเภทนี้ได้แก่เปอร์เซปตรอน (Perceptron) ,อะดาไลน์ (Adaline) , และการกระจายย้อนกลับ (Back-Propagation)

- **ข่ายงานนิเวศแบบไม่มีการชี้แนะ (Unsupervised Neural Network)**

ข่ายงานนิเวศที่มีการเรียนรู้แบบไม่มีการชี้แนะต้องการเพียงค่าอินพุตในการฝึกข่ายงานโดยไม่มีผู้สอนหรือค่าเป้าหมาย การปรับองค์ประกอบของข่ายงานนิเวศใช้สัญญาณเอาต์พุตอย่างเดียวในระหว่างการฝึกค่าน้ำหนักของข่ายงานจะถูกปรับโดยอินพุตที่คล้ายๆกันสร้างเอาต์พุตที่คล้ายกัน ซึ่งวิธีการปรับมีหลายวิธี ในแต่ละวิธีมีวัตถุประสงค์เพื่อให้การเรียนรู้ของข่ายงานประเภทนี้คล้ายกับสมองของมนุษย์ที่เปลี่ยนแปลงโครงสร้างของมันภายใต้อิทธิพลของประสบการณ์โดยไม่มีผู้สอน ตัวอย่างการเรียนรู้ประเภทนี้ได้แก่ AG (Additive Grossberg) และ Self Organizing Map (SOM) เป็นต้น



### 3.4.3 ข่ายงานนิวรัลแบบการกระจายย้อนกลับ

#### 3.4.3.1 มีโครงสร้างดังนี้

- การเชื่อมโยงแบบป้อนไปข้างหน้า (Feedforward Connection)
- ฟังก์ชันการเรียนรู้ที่มีการชี้้นำ (Learning Function) แบบการกระจายย้อนกลับ (Back propagation)
- ฟังก์ชันมูลฐานเชิงเส้น (Linear Basis Function)
- ฟังก์ชันกระตุ้น (Activation Function) เป็นฟังก์ชันซิกมอยด์
- ฟังก์ชันความผิดพลาด (Error Function) ใช้การออปติไมซ์ โดยนำเกรเดียนต์เดสเซนท์ (Gradient Descent) และ รีเวนเบอร์เกอร์ - มาร์ควอร์ดท์ (Levenberg – Marquardt) ซึ่งเป็นการลดค่าความผิดพลาดให้ต่ำสุด และทำการปรับค่าน้ำหนัก โดย Levenberg – Marquardt นั้นสามารถหาค่าตอบได้เร็วกว่า แบบ Gradient Descent มาก

#### 3.4.3.2 วิธีการออปติไมซ์ (Optimize Method)

กระบวนการที่ไม่เป็นเชิงเส้นหรือที่ซับซ้อนจะมีการเชื่อมโยงอินพุทไปสู่เอาต์พุทโดยใช้ชุดของน้ำหนักในข่ายงาน วัตถุประสงค์ของอัลกอริธึมก็คือต้องการหาค่าน้ำหนักให้ถูกต้องเพื่อที่สามารถจำลองแบบกระบวนการนั้นให้สามารถทำนายค่าเอาต์พุทจากอินพุทที่รับเข้าไปได้ถูกต้อง ดังนั้นอัลกอริธึมการกระจายย้อนกลับเป็นอัลกอริธึมหนึ่งที่ต้องการลดค่าผิดพลาดกำลังสองระหว่างค่าเป้าหมายกับผลลัพธ์จากข่ายงาน โดยใช้วิธีการออปติไมซ์แบบเกรเดียนต์เดสเซนท์ ซึ่งเป็นวิธีการลดความผิดพลาด (Optimization Method) ซึ่งแสดงในสมการ

$$\text{สำหรับข้อมูลแต่ละชุด} \quad \text{Minimize:} \quad E = \frac{1}{2} \sum_{i=1}^{N_L} (t_i - a_i(l))^2 \quad \text{สมการ 3.16}$$

$$\text{Subject to:} \quad \begin{aligned} t_{\max} &> t_i > t_{\min} && \text{สมการ 3.17} \\ a_{\max} &> a_i(l) > a_{\min} \end{aligned}$$

เมื่อ	$E$	คือค่าความผิดพลาด (Global Error) ที่ชั้นเอาต์พุท ได้มาจากผลรวมกำลังสองของความแตกต่างของเอาต์พุทที่ได้จากการคำนวณกับค่าเป้าหมายชั้นเอาต์พุท
	$t_i$	คือค่าเป้าหมาย
	$t_{\max}$	คือค่าสูงสุดของเป้าหมาย
	$t_{\min}$	คือค่าต่ำสุดของเป้าหมาย
	$a_i(l)$	คือค่าเอาต์พุทของข่ายงาน
	$a_{\max}$	คือค่าสูงสุดของเอาต์พุทของข่ายงาน
	$a_{\min}$	คือค่าต่ำสุดของเอาต์พุทของข่ายงาน



### 3.4.3.3 เกรเดียนต์เดสเซนท์(Gradient Descent : Cauchy,A.,1847,Luenberger,D.G.,1984)

เกรเดียนต์เดสเซนท์เป็นวิธีหนึ่งในการลดความผิดพลาดให้ต่ำที่สุดโดยการปรับค่าน้ำหนัก ดังนั้นถ้าทราบค่าอนุพันธ์ย่อย (Partial Derivative) ของค่าผิดพลาดเทียบกับค่าน้ำหนักแต่ละค่า จะทำให้รู้ทิศทางที่น้ำหนักไปในทิศทางที่ลดค่าผิดพลาด นั่นคือ  $\partial E/\partial w_{ij}$  เป็นค่าเกรเดียนต์ในทิศทางลบของพื้นผิวความผิดพลาด เครื่องหมายลบแทนด้วยทิศทางเกรเดียนต์ลดความผิดพลาด ดังนั้นวิธีการปรับค่าน้ำหนักใหม่ในทางลดความผิดพลาดให้ต่ำสุด ซึ่งควรจะเป็นความผิดพลาดจริง (Global Error) มากกว่าที่จะเป็นความผิดพลาดเฉพาะที่ (Local Error) สำหรับอัลกอริธึมการกระจายย้อนกลับจะทำการปรับค่าน้ำหนัก ( $w_{ij}$ ) เพื่อให้ค่า E มีค่าต่ำสุด สมการการเรียนรู้แบบเกรเดียนต์เดสเซนท์แสดงดังในสมการ

$$w_{ij}^{m+1}(l) = w_{ij}^m(l) + \Delta w_{ij}(l) \quad \text{สมการ 3.18}$$

อนุพันธ์ของค่าผิดพลาดจะใช้วิธีการตามเทคนิคของกฎลูกโซ่ (chain rule) ของชุดข้อมูลการฝึก (iteration)

$$\Delta w_{ij}^m(l) = -\eta \frac{\partial E}{\partial w_{ij}^m(l)} = -\eta \frac{\partial E}{\partial a_i^m(l)} \frac{\partial a_i^m(l)}{\partial w_{ij}^m(l)} \quad \text{สมการ 3.19}$$

และ พิจารณาส่วนที่ 1

$$\frac{\partial a_i^m(l)}{\partial w_{ij}^m(l)} = \frac{\partial f(u_i^m(l))}{\partial w_{ij}^m(l)} \quad \text{สมการ 3.20}$$

$$\begin{aligned} &= \frac{\partial f(\sum_{j=1}^N w_{ij}^m(l) a_j^m(l-1) + b_i(l))}{\partial w_{ij}^m(l)} \\ &= f'(u_i^m(l) a_j^m(l-1)) \end{aligned}$$

พิจารณาส่วนที่ 2

$$-\frac{\partial E}{\partial a_i^m(l)} = \delta_i^m(l) \quad \text{สมการ 3.21}$$

แทนค่ากลับจะได้

$$\Delta w_{ij}^m(l) = \eta \delta_i^m(l) f'(u_i^m(l) a_j^m(l-1)) \quad \text{สมการ 3.22}$$

$\eta$  คือ อัตราการเรียนรู้ (Learning rate) ถ้าเลือกอัตราการเรียนรู้ที่มีค่ามาก แต่ละขั้นตอนที่มีการปรับน้ำหนักค่าความผิดพลาดจะสูงเกินจุดต่ำสุด และนำไปสู่การแกว่งตัวรอบจุดต่ำสุด ในทางตรงกันข้ามถ้าอัตราการเรียนรู้มีค่าน้อยมากจะเป็นสาเหตุให้เข้าสู่จุดต่ำสุดช้าซึ่งต้องใช้ อัลกอริธึมที่ช่วยเร่งในการเรียนรู้เช่น Rprop หรือ Levenberg-Marquardt โดยที่ค่าน้ำหนักจะถูกปรับจนกระทั่งค่าความแตกต่างของค่าเป้าหมายกับผลลัพธ์จากข่ายงานมีค่าน้อยที่สุดในช่วงที่ยอมรับได้

### 3.4.3.4 ริเวนเบอร์เกอร์ - มาร์ควาร์ดท์ (K.Levenberg 1944 , D.Marquardt 1963)

การออปติไมซ์แบบเกรเดียนต์เดสเซนท์ เป็นการใช้อ่อนพจน์ย่อยอันดับหนึ่งในการหาจุดคำตอบและทำได้ดีสำหรับแบบจำลองอย่างง่าย (Simple Models) ที่มีความเป็นเชิงเส้นสูงเท่านั้น โดยกระบวนการเรียนรู้ของ เกรเดียนต์เดสเซนท์ นั้นจะกำหนดค่าถ่วงน้ำหนักเป็นขั้นๆและจะถูกปรับตำแหน่งใหม่ด้วยการเคลื่อนอย่างรวดเร็วไปยังทิศทางที่เรียกว่า Steepest Descent หลังจากปรับค่าถ่วงน้ำหนัก ความคาดเคลื่อนจะเลื่อนไปยังค่าต่างๆบนพื้นผิวความคาดเคลื่อน จนกระทั่งเคลื่อนเข้าสู่จุดคำตอบที่มีค่าต่ำที่สุด

ทิศทางการเคลื่อนตัวของความคาดเคลื่อนบนพื้นผิวขึ้นอยู่กับรูปร่างพื้นผิว ซึ่งก็คือรูปแบบของฟังก์ชันและค่าคงที่ในการเรียนรู้ บางพื้นผิวอาจมีลักษณะของพื้นราบ หรือแอ่งกระทะ ซึ่งค่าถ่วงน้ำหนักจะถูกปรับบ่อยครั้งเพื่อเคลื่อนไปหาความคาดเคลื่อนที่ต่ำกว่า ณ ตำแหน่งที่มีความคาดเคลื่อนสูง การใช้ขนาดขั้นของการปรับค่าถ่วงน้ำหนักมากอาจนำมาซึ่งความไม่เสถียรในการเคลื่อนตัวข้ามความชันพื้นผิว ในกรณีนี้จึงเป็นการยากสำหรับการออปติไมซ์แบบเกรเดียนต์เดสเซนท์ที่จะเลือกทิศทางที่ถูกต้องและต้องเคลื่อนตัวอย่างช้าๆ และไร้ทิศทาง Jacob(1988) พบว่าเส้นความชันของพื้นผิวความคาดเคลื่อนจะเป็นวงรี (Elliptic) เนื่องจากพื้นผิวความคาดเคลื่อนเป็นการรวมกันของเทอมยกกำลังสอง (Quadratic Term) ดังนั้นเกรเดียนต์เดสเซนท์จะทำให้ทิศทางการเคลื่อนตัวของความคาดเคลื่อนไม่เข้าสู่ค่าต่ำสุด (เลรี ศุภราทิพย์) ดังนั้นฟังก์ชันของแบบจำลองที่มีความซับซ้อนสูง (Complex Models) และความไม่เป็นเชิงเส้นสูง แล้วนั้นยากต่อการหาจุดคำตอบ อีกทั้งการลู่เข้าสู่จุดคำตอบก็ใช้เวลานาน ปัญหาดังกล่าว Levenberg และ Marquardt ได้ออกแบบโดยนำอนุพันธ์ย่อยอันดับสอง มาใช้ในการแก้ปัญหาเกี่ยวกับแบบจำลองที่มีความซับซ้อนสูง โดยการปรับปรุง และพัฒนาอัลกอริธึมของการออปติไมซ์แบบเกรเดียนต์เดสเซนท์ ซึ่งในกรณีที่ ค่าฟังก์ชันความคาดเคลื่อน  $E(w)$  นั้นมีความไม่เป็นเชิงเส้นสูงจึงทำการ Linearize ฟังก์ชันนี้ เพื่อนำไปประมาณค่าความคาดเคลื่อนแบบเป็นเชิงเส้นมากขึ้น และนำไปหาจุดคำตอบของค่าความคาดเคลื่อน เริ่มแรกเราพิจารณาฟังก์ชันค่าความคาดเคลื่อนที่ได้จากการเรียนรู้ของข่ายงาน เพื่อนำมาออปติไมซ์เพื่อหาค่าถ่วงน้ำหนักเพื่อลดฟังก์ชันค่าความคาดเคลื่อน โดยเราจะทำการ Linearize จะได้ฟังก์ชันใหม่ดังนี้

$$\hat{f}(x; w) = f(x; w_0) + (w - w_0)^T \nabla f(x; w_0) \quad \text{สมการ 3.23}$$

โดย  $f(x; w_0)$  สมมติให้เป็นฟังก์ชันค่าความคาดเคลื่อนแบบไม่เป็นเชิงเส้น

ส่วน  $\hat{f}(x; w_0)$  การประมาณฟังก์ชันค่าความคาดเคลื่อนแบบเชิงเส้น

และเราจะนำฟังก์ชันนี้ใช้การหา The quadratic approximation ได้ดังนี้ โดย รูปแบบอย่างง่ายของ  $\nabla \hat{f}(x; w)$ :  $\nabla \hat{f}(x; w) = \nabla f(x; w)$  เนื่องจาก  $\hat{f}$  เป็นฟังก์ชันเชิงเส้นของ  $w$



$$\hat{E}(w) = \left\langle \left( \hat{f}(x; w) - y \right)^2 \right\rangle \quad \text{สมการ 3.24}$$

$$\nabla \hat{E}(w) = \left\langle 2(f(x, w) - y) \nabla f(x, w) \right\rangle$$

$$\nabla \hat{E}(w) = \left\langle 2(f(x; w_0) + (w - w_0)^T \nabla f(x; w_0) - y) \nabla f(x; w_0) \right\rangle$$

และ

$$\nabla \hat{E}(w) = \left\langle 2((f(x; w_0) - y) \nabla f(x; w_0) + (w - w_0)^T \nabla f(x; w_0)^T \nabla f(x; w_0)) \right\rangle \quad \text{สมการ 3.25}$$

$d = (f(x; w_0) - y) \nabla f(x; w_0)$  คือ อนุพันธ์ย่อยอันดับหนึ่ง เรียกว่า *derivative*

$H = \nabla f(x; w_0)^T \nabla f(x; w_0)$  คือ อนุพันธ์ย่อยอันดับสอง เรียกว่า *Hessian*

จะได้ว่า

$$\nabla \hat{E}(w) = 2d + 2(w - w_0)H \quad \text{สมการ 3.26}$$

$$\nabla \hat{E}(w) = 0; 2(w - w_0)H + 2d = 0$$

$$w_{opt} = -H^{-1}d + w_0$$

ซึ่งคำตอบนี้ *Levenberg* ใช้ความสัมพันธ์ทั้ง 2 แบบมารวมกันได้ดังนี้

$$1. w_{i+1} = w_i - \mu d \quad \text{ที่แสดงถึงวิธี } \textit{steepest descent}$$

$$2. w_{i+1} = w_i - H^{-1}d \quad \text{ที่แสดงถึงวิธี } \textit{The quadratic approximation}$$

โดยหัวใจของเทคนิคนี้คือการนำข้อดีทั้งสองมารวมกันระหว่าง *steepest descent* ซึ่ง

เป็นส่วนช่วยในการหาค่าคำตอบ ส่วน *The quadratic approximation* ทำหน้าที่ในส่วนการประมาณฟังก์ชันค่าความคาดเคลื่อนที่มีความไม่เป็นเชิงเส้นสูง อัลกอริทึมของ *Levenberg* คือ

$$w_{i+1} = w_i - (H + \lambda I)^{-1}d \quad \text{ที่แสดงถึงวิธี } \textit{Levenberg's method}$$

โดย  $\lambda$  คือ bending factor สามารถแบ่งเป็นส่วนของการออฟติไมซ์เพื่อหาค่าถ่วงน้ำหนักได้ดังนี้

เมื่อ  $\lambda \ll$  มีค่าน้อยสมการนี้ก็จะ เป็นแบบ *The quadratic approximation*

$\lambda \gg$  มีค่ามากสมการนี้จะเป็นแบบ *steepest descent*

ซึ่งเป็นที่รู้กันว่า เมื่อค่าความคาดเคลื่อนเพิ่มขึ้น *The quadratic approximation* ไม่สามารถ

ใช้งานได้ดี และไม่สามารถเข้าสู่จุดต่ำสุดได้ ดังนั้นจึงต้องเพิ่ม  $\lambda$  เพื่อให้สมการเป็นแบบ

*steepest descent* มากขึ้น ในทางกลับกัน ถ้าค่าความคาดเคลื่อนลดลง การประมาณจะ

สามารถใช้งานได้ดี และคาดว่าจะกำลังจะเข้าสู่จุดคำตอบที่ต่ำสุดแล้ว ดังนั้นจึงทำการลดค่า  $\lambda$

เพื่อให้สมการเป็นแบบการประมาณมากขึ้น ต่อมา *Marquardt* ลดข้อจำกัดเหล่านี้ ของ

*Levenberg* กล่าวคือขณะที่  $\lambda$  มากสมการจะมีความเป็นแบบ *steepest descent* สูง และเพื่อ

ต้องการนำข้อดีของการประมาณโดย *Hessian Matrix* โดย *Marquardt* แทน

*Identity matrix* ด้วย *diagonal Hessian matrix* ดังนี้

$$w_{i+1} = w_i - (H + \lambda \text{diag}[H])^{-1} d \quad \text{สมการ 3.27}$$

ผลลัพธ์คือเป็นสมการการปรับค่าถ่วงน้ำหนักของข่ายงานนิวรัลที่มีความนิยมใช้มากที่สุดเนื่องจากเป็นการรวมข้อดีทั้งหมดของการออฟติไมซ์ซึ่งความเร็วจะสูงกว่า เกรเดียนต์เดสเซนท์ ที่มีค่า  $\eta$  อัตราการเรียนรู้ (Learning rate) มาก



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย