ECH4905 - Special Topics in ChemE - Optimization

Andres Espinosa

February 13, 2025

Contents

1 Chapter 1 - Optimization Models

1.1 Tuesday 01/14/2025

1.1.1 Modeling Overview

An optimization problem is represented by degrees of freedom that inform decisions. The degrees of freedom of a problem dictate the number of decisions that can be made.

- n Decisions $> 0 \implies$ optimization problem
- n decisions $= 0 \implies$ simulation

The number of decisions is equal to the degrees of freedom, which is equivalent to the number of variables minus the number of inequalities.

1.1.2 Example Problem

This example problem will use a toy example of minimizing the perimeter of a rectangle with side lengths x and y.

$$\begin{aligned} \text{minimize} \quad & P && (1)\\ \text{subject to} \quad & xy = 2000 && (2)\\ & P = 2x + 2y && (3)\\ & x \geq 0, y \geq 0 && (4) \end{aligned}$$

with variables x, y, P. Since there are two equalities in the optimization problem above and three variables, there is one degree of freedom. This can be seen more plainly by reformulating the problem into the below problem

$$\begin{aligned} \text{minimize} \quad & 2(\frac{2000}{y}) + 2y && (5)\\ \text{subject to} \quad & y \geq 0 && (6) \end{aligned}$$

1.1.3 General Optimization Definitions

A general optimization problem can be written as

$$\begin{aligned} \text{minimize} \quad & F(x, y) && (7)\\ \text{subject to} \quad & h(x, y) = 0 && (8)\\ & g(x, y) \leq 0 && (9) \end{aligned}$$

where $x \in \mathbb{R}^n, y \in \{0, 1\}^m$. This general optimization problem can be broken down into different problems with levels of complexity.

Firstly, a **linear program** (LP) is defined as

$$\begin{aligned} \text{minimize} \quad & F(x) && (10)\\ \text{subject to} \quad & h(x) = 0 && (11)\\ & g(x) \leq 0 && (12) \end{aligned}$$

Where $x \in \mathbb{R}^n$ and F, h, g are affine functions. Linear problems typically make many approximations about the real physics/constraints that are in the world. These approximations are a tradeoff with computational tractability. The Simplex method is typically used to solve linear programs.

Secondly, a **non-linear program** (NLP) is defined as

$$\begin{aligned} \text{minimize} \quad & F(x) && (13)\\ \text{subject to} \quad & h(x) = 0 && (14)\\ & g(x) \leq 0 && (15) \end{aligned}$$

Where $x \in \mathbb{R}^n$ and any of the functions F, h, g are non-linear. A subset of NLPs are those problems that are *convex*. Convexity in a non-linear program is achieved when F, g are convex functions and h is an affine function. Some algorithms that are used to solve NLPs are SQP, IPOPT.

A higher degree of complexity, are **mixed integer linear programs** (MILP) defined as

$$\begin{aligned} \text{minimize} \quad & F(x, y) && (16)\\ \text{subject to} \quad & h(x, y) = 0 && (17)\\ & g(x, y) \leq 0 && (18) \end{aligned}$$

where $x \in \mathbb{R}^n, y \in \{0, 1\}^m$. and F, h, g are affine functions which can be represented as $F(x, y) = Ax + By$. These problems have applications in chemical engineering with examples such as biomass, waste, and fuel supply chains. In a biomass biorefinery location problem, you would have a set of locations to pick from y and you want to minimize the cost. An algorithm used to solve these problems is branch-and-bound.

With the highest degree of complexity, a **mixed integer non-linear program** (MINLP) is defined as

$$\begin{aligned} \text{minimize} \quad & F(x, y) && (19)\\ \text{subject to} \quad & h(x, y) = 0 && (20)\\ & g(x, y) \leq 0 && (21) \end{aligned}$$

where $x \in \mathbb{R}^n, y \in \{0, 1\}^m$ and any of the F, h, g functions are non-linear. BARON, SCIP, MAINGO are algorithms that have been developed in the last 10 years in response to solve MINLPs. Chemical engineering is naturally non-linear and there has been a big push from the chemical optimization community to create solutions for MINLPs. These are complex problems but also important problems.

1.1.4 Specific Model Criteria

Optimization problems can be also defined by the **number of objectives** they have.

- Single objective
- Multiple objective

Multiple objectives introduce a tradeoff between different objectives with each other. If the number of objectives is low, then it is easy to identify and visualize the tradeoffs. The concept of pareto optimal will be explored further in the course.

$$\text{minimize} \quad \langle F_1(x), F_2(x), \ldots, F_k(x) \rangle \tag{22}$$

Optimization problems can also have a level of **uncertainty** in the problem. For example, the price and cost parameters may be random and uncertain. The majority of time, the randomness and uncertainty in models are ignored by modelers. A good guess is used and substituted for the random parameters such as an average or other estimate. However, it is possible to create stochastic models that attempt to handle uncertainty.

- Deterministic models: Uncertainty is ignored and stochastic parameters are replaced by a fixed "good guess".
- Stochastic models: Uncertainty is handled and incorporated in the constraints.

Typically, stochastic models are significantly more difficult to compute than deterministic models. So, a deterministic model should be used and solved before a stochastic one is approached.

Optimization models are also specified by their **type/number of constraints**

- Unconstrained
- Constrained

1.1.5 Optimization Concepts

For general optimization problem defined above, a **feasible region** is defined as the set of $\langle x, y \rangle$ that satisfy all equalities and inequalities.

$$\begin{aligned}
\text{minimize} \quad & -xy && (23)\\
\text{subject to} \quad & x \geq 0, x \leq 1 && (24)\\
& y \geq 0, y \leq 1 && (25)
\end{aligned}$$

For this problem, the feasible region R is

$$R(x, y) = \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1\} \tag{26}$$

A **local minimum** with respect to a distance ε is defined as a point (x^*, y^*) that satisfies $F(x^*, y^*) \leq F(x, y), \|x^* - x\| \leq \varepsilon$. A **global minimum** is a point that satisfies $F(x^*, y^*) \leq F(x, y)$.

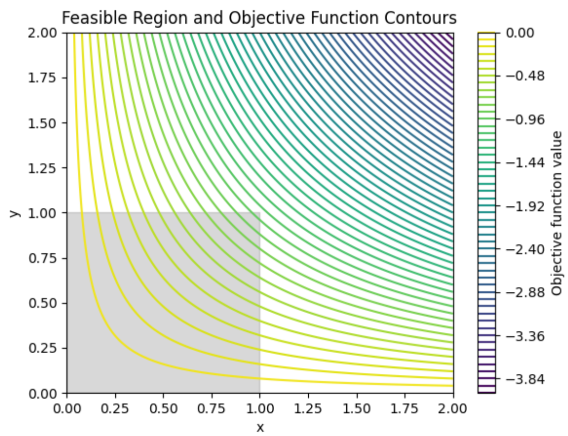


Figure 1: A plot of the optimization problem $F(x, y) = -xy$

2 Chapter 2 - Mathematics Review

2.1 Thursday 01/16/2025

2.1.1 Chapter Math Outline

The chapters over the next few lessons will include some math that will be used in different chapters.

- Chapter 1: Linear algebra
  - Matrices
  - Eigenvalues
- Chapter 2: Convex Analysis
  - Convexity
  - Quadratic forms
  - Taylor Series

2.1.2 Linear Algebra Review

A vector $\mathbf{x} \in \mathbb{R}^n$ is defined as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{27}$$

This vector represents a direction and magnitude in n dimensions.
The $l2$-norm of a vector is defned as

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots x_n^2} \tag{28}$$

Vector addition is defined as

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \tag{29}$$

Scalar multiplication of a scalar $a \in \mathbb{R}$ with a vector $\mathbf{x} \in \mathbb{R}^n$

$$a\mathbf{x} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix} \tag{30}$$

The dot product of two vectors $\mathbf{x}, \mathbf{y}, \in \mathbb{R}^n$ is defined as

$$\mathbf{x}^\top \mathbf{y} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{31}$$

With equivalent notation $\mathbf{x}^\top \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$

$$\|\mathbf{x} - \mathbf{y}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 - 2\|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \cos\theta \tag{32}$$

$$\|\mathbf{x} - \mathbf{y}\|_2^2 = (\mathbf{x} - \mathbf{y})^\top (\mathbf{x} - \mathbf{y}) \tag{33}$$

The Cauchy-Schwarts Inequality is derived by the following

$$\cos\theta = \frac{x^\top y}{\|x\|\|y\|} \tag{34}$$

$$\frac{x^\top y}{\|x\|\|y\|} \leq 1 \tag{35}$$

$$x^\top y \leq \|x\|\|y\| \tag{36}$$

The triangle inequality is defined as

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \tag{37}$$

A matrix $A \in \mathbb{R}^{m \times n}$ is defined as

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \tag{38}$$

The transpose of a matrix $A^\top$ flips each value for the row and column as such

$$A^\top = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \tag{39}$$

Matrix addition is element-wise and can be shown as such between two matrices $A, B \in \mathbb{R}^{m \times n}$

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{bmatrix} \tag{40}$$

Scalar multiplication of a matrix A with a scalar α is defined as

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \dots & \alpha a_{1n} \\ \vdots & \dots & \vdots \\ \alpha a_{m1} & \dots & \alpha a_{mn} \end{bmatrix} \tag{41}$$

Matrix multiplication between two matrices $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$ is defined as

$$AB = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1p} \\ \vdots & \dots & \vdots \\ b_{n1} & \dots & b_{np} \end{bmatrix} \tag{42}$$

With the following properties

- $AB \neq BA$
- $(ABC)^\top = C^\top B^\top A^\top$

The inverse of a matrix has the following properties

- $AA^{-1} = I$
- $(AB)^{-1} - B^{-1}A^{-1}$
- $(B^\top)^{-1} = (B^{-1})^\top$
- $(A^{-1})^{-1} = A$

Orthonormal matrices have the properties

- $O^\top O = I$
- $O^\top = O^{-1}$

Some following matrix partitions are useful

$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} + \begin{bmatrix} A_2 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_1 + A_1 \\ B_1 + B_2 \end{bmatrix} \tag{43}$$

$$\begin{bmatrix} A \\ B \end{bmatrix}^\top = \begin{bmatrix} A^\top & B^\top \end{bmatrix} \tag{44}$$

$$A \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}^\top = \begin{bmatrix} AB_1 & AB_2 \end{bmatrix} \tag{45}$$

The determinant of a matrix $A \in \mathbb{R}^{n \times n}$ is calculated by subtracting the product of the diagonals of a matrix recursively. The minor of a matrix is the section of a matrix that is achieved when removing a section. Simply, the determinant of a matrix can also be defined as the product of the eigenvalues.

$$\det A = \prod \lambda \tag{46}$$

2.1.3 Spaces

A vector space is defined as a set of all vectors with some properties. We define a vector space $\mathbb{V}$

$$u + v \in \mathbb{V} \tag{47}$$
$$u + (v + w) = (u + v) + w \tag{48}$$
$$u + v = v + u \tag{49}$$

A linearly dependent system of vectors $v_i \in \mathbb{R}^n$ satifies the following for a set of scalars c_i

$$c_1 v_1 + \cdots + c_m v_m = 0 \tag{50}$$

The span of a set of vectors v_i is the set of all vectors that can be created with a linear combination of those vectors.

$$\{x | c_1 v_1 + \cdots + c_m v_m = x\} \tag{51}$$

The basis of a space is the minimum number of vectors needed to span a vector space.

2.2 Tuesday 01/21/2025

2.2.1 Matrix eigenvalues

We define λ as the eigenvalues of a matrix. We investigate what it means for matrices with different eigenvalues.

- $\lambda_i \geq 0 \rightarrow$ Matrix A is positive semi-definite
- $\lambda_i > 0 \rightarrow$ Matrix A is positive definite
- $\lambda_i \leq 0 \rightarrow$ Matrix A is negative semi-definite
- $\lambda_i < 0 \rightarrow$ Matrix A is negative definite
- $\lambda_i < 0, \lambda_i > 0 \rightarrow$ Matrix A is indefinite

2.2.2 Convex Sets

A convex set is defined as a set where the line through all points is contained in the same set.

Half Spaces:
Half spaces are the space under a line, defined with the parameters $c \in \mathbb{R}^n z \in \mathbb{R}$,

$$\{x|c^\top x \leq z, x \in \mathbb{R}^n\} \tag{52}$$

A half space is open if the inequality is strict.
The intersection of a finite number of closed half spaces is known as a polytope. If a polytope is bounded on all sides, it is also known as a polyhedron.

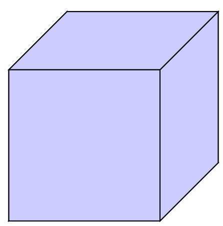


Figure 2: A polyhedron

Extreme Points:
Consider a convex set $C \subseteq \mathbb{R}^n$. Let $z \in \mathbb{R}^n$, z is an extreme point of C if $z \in C$ and there are no $x, y \in C$ such that $z = (1-\lambda)x + \lambda y$. In other words, the point z is not on a line defined by two points. The figure 3 shows a polyhedron. It is impossible to represent the red points as a positive combination of two other points in this square. Therefore, they are extreme points.

Convex combinations:
Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be a set of m vectors in $\mathbb{R}^n$. A convex combination of these vectors is a point of the form

$$\lambda_1 \mathbf{x}_1 + \cdots + \lambda_m \mathbf{x}_m \tag{53}$$

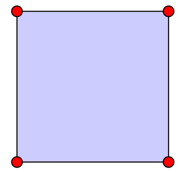


Figure 3: A 2D polyhedron with emphasized vertices

where $\mathbf{1}^\top \lambda = 1, \lambda \succeq 0$

A Simplex:
A simplex is the simplest possible polytope in a given dimension. Figure 4 shows the simplexes in 0 to 3 dimensions.

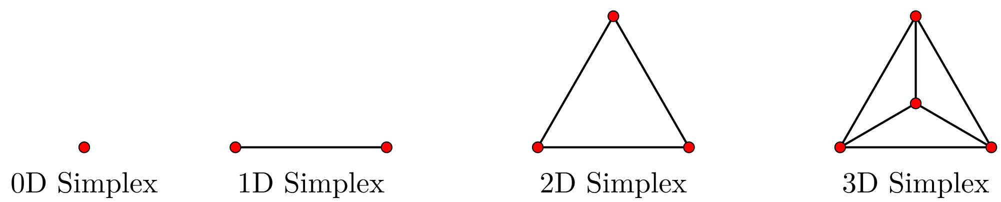


Figure 4: Simplexes in 0D to 3D

Convex Hull
The convex hull of a set S is defined as the smallest convex set that contains S. It is also defined as the intersection of all convex sets which contain S. Figure 5 shows the convex hull of a nonconvex set.

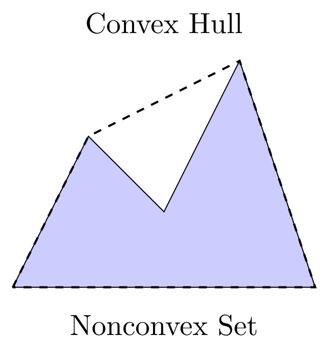


Figure 5: A 2D nonconvex set and its convex hull

Farkas Theorem:
The Farkas Theorem is a theorem that allows for a certificate of feasibility (or infeasibility) of an optimization problem. With the variables $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^n$, exactly one of the following two propositions is true.

- $Ax = b, x \succeq 0$ for some $x \in \mathbb{R}^n$. In other words, b is in the cone spanned by convex combinations of the columns of A.

- $A^\top y \succeq 0, b^\top y \leq 0$, for some $y \in \mathbb{R}^m$. In other words, the angle between y and b is greater than 90 and the angle between y and every column of A is less than 90.

$$Ax = b \rightarrow \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{54}$$

2.2.3 Convex Functions

Derivatives:
Gradient refresher

$$\nabla f(x) = \begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \\ \vdots \\ \frac{df}{dx_n} \end{bmatrix} \tag{55}$$

A directional derivative is the transpose of the gradient times a direction $\nabla f(x)^\top d$. Consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ that is differentiable. A direction is a descent direction if $d^\top \nabla f(x) < 0$. The jacobian of a function is a generalized derivative for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. It is defined as

$$J(x) = \begin{bmatrix} \nabla f_1(x)^\top \\ \nabla f_2(x)^\top \\ \vdots \\ \nabla f_m(x)^\top \end{bmatrix} \tag{56}$$

The hessian of a function is the second derivative of a function. The Hessian is symmetric as long as the mixed derivatives are equal to each other.

$$H(f(x)) = \begin{bmatrix} \frac{d^2 f}{dx_1^2} & \dots & \frac{d^2 f}{dx_1 dx_n} \\ \vdots & \vdots & \vdots \\ \frac{d^2 f}{dx_n dx_1} & \dots & \frac{d^2 f}{dx_n^2} \end{bmatrix} \tag{57}$$

Convex Functions:
Let C be a convex subset of $\mathbb{R}^n$, and f(x) be a real-valued function defined on C. The function f is convex if Jensens inequality below holds. A function is strictly convex if the below property holds with strict inequality.

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \tag{58}$$

In words, a function is convex if the line between any two points lies above every point of the function between those two points. Figure 6 shows a convex function and how the line between any two points lies above the graph. Another way to define convex functions is with respect to the Hessian of the function $H(f(x))$. A function is convex if the Hessian of the function is positive semi definite $H(f(x)) \succeq 0$. Convexity can also be defined in sections of a function. For example, let S

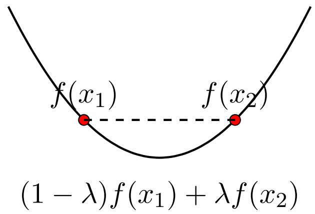


Figure 6: A convex function and the line between two points

be anon-empty open set in $\mathbb{R}^n$. If $f(x_0)$ is convex at the point x_0, then H is positive semi-definite at that point.

Example Problem:
Determine if the following equation is convex

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 \tag{59}$$

$$\nabla f(x) = \begin{bmatrix} 2x_1 + 2x_2 \\ 2x_2 + 2x_1 \\ 2x_3 \end{bmatrix} \tag{60}$$

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{61}$$

Calculating the eigenvalues

$$\det \begin{bmatrix} 2-\lambda & 2 & 0 \\ 2 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{bmatrix} \tag{62}$$

$$(2-\lambda)(2-\lambda)(2-\lambda) - 2(2)(2-\lambda) + 0 = 0 \tag{63}$$
$$(2-\lambda)^3 - 4(2-\lambda) = 0 \tag{64}$$
$$(2-\lambda)((2-\lambda)^2 - 4) = 0 \tag{65}$$
$$\lambda = [2, 0, 4] \tag{66}$$

$\lambda = [0, 2, 4] \succeq 0$, therefore the hessian is positive semi-definite and the function is convex.

2.2.4 Properties of Convex Functions

- if f_i are convex, $\sum_i f_i(x)$ is convex
- if f is convex, $\lambda f(x)$ is convex, where $\lambda isascalar$
- Let f be convex, and g be an increasing function. The convex function $g(f(x))$ is also convex.

2.2.5 Quadratic Forms

The quadratic form of a vector $x \in \mathbb{R}^n$ with parameters $Bin\mathbb{R}^{n \times n}, a \in \mathbb{R}^n, c \in \mathbb{R}$ is

$$f(x) = \frac{1}{2}x^\top Bx + a^\top x + c \tag{67}$$

2.3 Thursday 01/23/2025

2.3.1 Eigenvalues and Eigenvectors

A non-singular matrix is a matrix A that follows the property $\det A \neq 0$.

Eigenvalues:
In order to calculate the eigenvalues of a matrix A, we solve the equation $\det(A - \lambda I) = 0$. For example, we take the matrix

$$A = \begin{bmatrix} 13 & -4 \\ -4 & 7 \end{bmatrix} \tag{68}$$

We subtract λ from the diagonals to get

$$A = \begin{bmatrix} 13 - \lambda & -4 \\ -4 & 7 - \lambda \end{bmatrix} \tag{69}$$

$$= (13 - \lambda)(7 - \lambda) - 16 = 0 \tag{70}$$

$$\lambda_1 = 15, \lambda_2 = 5 \tag{71}$$

Not all eigenvalues are always real. The number of eigenvalues is the same as the dimension of the matrix but is not always real.

Eigenvectors
To find the eigenvectors ν of the matrix, we solve $(A - \lambda I)\nu = 0$.

$$A - \lambda I = (\begin{bmatrix} 13 & -4 \\ -4 & 7 \end{bmatrix} - 15 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{72}$$

$$= \begin{bmatrix} -2 & -4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{73}$$

$$-2v_1 - 4v_2 = 0 \tag{74}$$

$$v_1 + 2v_2 = 0 \tag{75}$$

$$v_1 = -2v_2 \tag{76}$$

$$\text{Infinite solutions, as long as the vector looks like } \begin{bmatrix} -2a \\ a \end{bmatrix} \tag{77}$$

Now we use the other eigenvalue $\lambda = 5$

$$A - \lambda I = (\begin{bmatrix} 13 & -4 \\ -4 & 7 \end{bmatrix} - 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}) \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{78}$$

$$= \begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{79}$$

$$\text{Infinite solutions, as long as the vector looks like } \begin{bmatrix} a/2 \\ a \end{bmatrix} \tag{80}$$

2.3.2 Convexity in Functions

A bilinear function is a function in the form

$$f(x_1, x_2) = x_1 x_2 \tag{81}$$

The hessian of this function looks like

$$H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{82}$$

The eigenvalues of this matrix are $(-\lambda)^2 - 1 = 0, \lambda = \pm 1$. This matrix is indefinite and therefore non-convex. This function creates a saddle and is non-convex in both variables but is convex in either variable at a fixed falue of the other.

3 Chapter 3 - Linear Programming

3.1 Tuesday 01/28/2025

3.1.1 Linear Tricks

Constraint Relaxation: If we have the following optimization problem

$$\begin{aligned} \text{minimize} \quad & c^\top x && (83) \\ \text{subject to} \quad & \sum_j a_j x_j \leq b && (84) \end{aligned}$$

, we can relax the constraint by adding a variable u that will allow us to bend the constraint a bit at a cost.

$$\begin{aligned} \text{minimize} \quad & c^\top x + cost(u) && (85) \\ \text{subject to} \quad & \sum_j a_j x_j - u \leq b && (86) \end{aligned}$$

Handling Absolute Values: Absolute values can be handled by introducing additional variables and constraints. One way is to split the variable into a positive and negative component.

$$\begin{aligned} \text{minimize} \quad & c^\top |x| && (87) \\ \text{subject to} \quad & Ax \leq b && (88) \end{aligned}$$

Here, we replace $|x|$ with $x^+ - x^-$

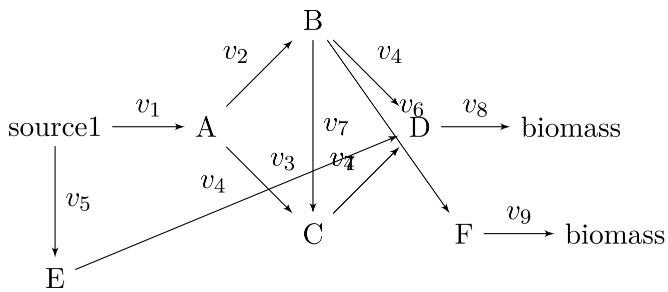
$$\begin{aligned} \text{minimize} \quad & c^\top (x^+ - x^-) && (89) \\ \text{subject to} \quad & Ax \leq b && (90) \\ & x = x^+ - x^- && (91) \\ & x, x^+, x^- \geq 0 && (92) \end{aligned}$$

3.1.2 Example Models

Metabolic Flux Balance: A metabolic flux balance problem is a problem where we have a cell that inputs some nutrient and outputs a metabolic yield, the growth of the cell. We have a *source* that is an input into a cell A at a rate v_1. A inputs into B at a rate v_2. In parallel, A inputs into C at a rate v_3. B + C input into $2D$ at a rate v_4, etc, the rest of the reaction network model is shown below.

- $source \to A \quad (v_1)$
- $A \to B \quad (v_2)$
- $A \to C \quad (v_3)$
- $B + E \to 2D \quad (v_4)$
- $source \to E \quad (v_5)$
- $B \to C + F \quad (v_6)$

- $C \to D \quad (v_7)$
- $D \to biomass \quad (v_8)$
- $F \to biomass \quad (v_9)$



In this problem, we have an objective we could have multiple goals, one is to maximize the weighted biomass output $w_8v_8+w_9v_9$. We have constraints that must obey the laws of chemistry and physics. One of them is *mass balance* constraints. This constraint ensures that the rate of consumption of each edge entering a node is equal to the sum of the rate of production leaving the node. Flow conservation constraints:

$$\begin{aligned} \text{maximize} \quad & w_8v_8 + w_9v_9 && (93)\\ \text{subject to} \quad & v_1 = v_2 + v_3 \quad (A) && (94)\\ & v_2 = v_4 + v_6 \quad (B) && (95)\\ & v_3 + v_6 = v_7 \quad (C) && (96)\\ & 2v_4 + v_7 = v_8 \quad (D) && (97)\\ & v_5 = v_4 \quad (E) && (98)\\ & v_6 = v_9 \quad (F) && (99) \end{aligned}$$

This problem will be unbounded if we do not constraint the entering sources, so we also introduce *source* constraints.

Oil blending Consider the problem where we have the following oils with different prices and hardness ratings for each

- veg 1 - \$110 - 8.8
- veg 2 - \$120 - 6.1
- oil 1 - \$130 - 2
- oil 2 - \$110 - 4.
- oil 3 - \$115 - 5

We sell each quantity for \$150. We have 200 tons of vegetable we can make oil with and 250 tons of non-vegetables we can make oil with. We want a hardness level between 3 and 6. To solve this, we can take the vector $\mathbf{x} \in \mathbb{R}^5$ as the amount of each liquid i we use. We define $\mathbf{x}_v$ and $\mathbf{x}_n$ subscripts as the subsets of $\mathbf{x}$ that are vegetable and non-vegetable. We similarly define $\mathbf{c}$ as the cost vector and $\mathbf{h}$ as the hardness vector. The selling price is $p = 150$. This problem can be formulated as

$$
\begin{aligned}
\text{maximize} \quad & 150(\mathbf{1}^\top \mathbf{x}) - \mathbf{c}^\top \mathbf{x} && (100)\\
\text{subject to} \quad & \mathbf{1}^\top \mathbf{x}_v \leq 200 && (101)\\
& \mathbf{1}^\top \mathbf{x}_n \leq 250 && (102)\\
& 3(\mathbf{1}^\top \mathbf{x}) \leq \mathbf{h}^\top \mathbf{x} \leq 6(\mathbf{1}^\top \mathbf{x}) && (103)\\
& \mathbf{x} \succeq 0 && (104)
\end{aligned}
$$

3.2 Thursday 01/30/2025

3.2.1 Example Models Continued

Oil Blending Over Time We now take the oil blending example previously and now consider it over time. We can buy different raw oil components at different periods of time and store them for use at a later time. Table 1 shows the raw oil prices over the next four months

	V1	V2	O1	O2	O3
M1	110	120	130	110	115
M2	130	130	110	90	115
M3	120	110	120	120	125
M4	90	100	140	80	135

Table 1: Oil prices over time

Our storage constraints are that we can store up to 1000 tons of each raw oil each period. We start with 500 tons of each oil and must end the total time with 500 tons of each oil. Going into each period, there is a \$5 cost for holding each ton into the next period. Our quality constraints are the same, that the total level of hardness must be between 3 and 6. Table 2 shows the hardness rating of each oil. Our capacity constraints are that we can not process more than 200 tons of vegetable oil and 250 non-vegetable oil per period.

Oil	Hardness
V1	8.8
V2	6.1
O1	2.0
O2	4.2
O3	5.0

Table 2: Oil hardness ratings

In order to tackle this problem, we can define the matrix $\Lambda \in \mathbb{R}^{5\times 5}$ as the cost matrix shown in table 1. We pad the first row with a 0 vector to represent us starting the first period with inventory. The matrix $\mathbf{X} \in \mathbb{R}^{5\times 5}$ is our decision variable where each row corresponds to the amount of each oil at the end of each period. The first and last column must be equal to a vector of 500s, since we must start and end with 500 of each raw oil. We also introduce buying and selling variables $\mathbf{B} \in \mathbb{R}^{5\times 5}, \mathbf{S} \in \mathbb{R}^{5\times 5}$ where each row is the amount of an oil we buy and sell before the end of each period. The first column is $\mathbf{0}$ since we can not buy or sell in the 0th period. The vector $\mathbf{h} \in \mathbb{R}^5$ are the hardness ratings per raw oil.

The buying cost component of the objective function is equal to $\textbf{trace}(\Lambda\mathbf{X})$. The inventory holding cost component of the objective function is equal to $5(\mathbf{1}^\top\mathbf{X}\mathbf{1} - 500 \times 5)$, the -500×5 term is to address the fact that we pay no inventory cost for the 0th period. The selling component of the objective function is equal to $150(\mathbf{1}^\top\mathbf{S}\mathbf{1})$

$$\Lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 110 & 120 & 130 & 110 & 115 \\ 130 & 130 & 110 & 90 & 115 \\ 120 & 110 & 120 & 120 & 125 \\ 90 & 100 & 140 & 80 & 135 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 500 & x_{11} & x_{12} & x_{13} & 500 \\ 500 & x_{21} & x_{22} & x_{23} & 500 \\ 500 & x_{31} & x_{32} & x_{33} & 500 \\ 500 & x_{41} & x_{42} & x_{43} & 500 \\ 500 & x_{51} & x_{52} & x_{53} & 500 \end{bmatrix} \tag{105}$$

$$\mathbf{B} = \begin{bmatrix} 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{31} & b_{32} & b_{33} & b_{34} \\ 0 & b_{41} & b_{42} & b_{43} & b_{44} \\ 0 & b_{51} & b_{52} & b_{53} & b_{54} \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 0 & s_{11} & s_{12} & s_{13} & s_{14} \\ 0 & s_{21} & s_{22} & s_{23} & s_{24} \\ 0 & s_{31} & s_{32} & s_{33} & s_{34} \\ 0 & s_{41} & s_{42} & s_{43} & s_{44} \\ 0 & s_{51} & s_{52} & s_{53} & s_{54} \end{bmatrix} \tag{106}$$

The objective function is therefore

$$\text{maximize} \quad 150(\mathbf{1}^\top\mathbf{S}\mathbf{1}) - \textbf{trace}(\Lambda\mathbf{X}) - 5(\mathbf{1}^\top\mathbf{X}\mathbf{1} - 500 \times 5) \tag{107}$$

The storage constraints can be expressed as $\mathbf{X}_{o,t} \leq 1,000 \quad \forall o \in O, t \in T$. Where O is the set of raw oils and T is the set of periods.
The capacity constraints can be modeled as $b_{1t}+b_{2t} \leq 200 \quad \forall t \in T$ and $b_{3t}+b_{4t}+b_{5t} \leq 250 \quad \forall t \in T$.
The hardness constraints can be modeled as $\mathbf{S}^\top\mathbf{h} \preceq 6(\mathbf{S}^\top\mathbf{1}), \mathbf{S}^\top\mathbf{h} \succeq 3(\mathbf{S}^\top\mathbf{1})$
We also must model our logical inventory constraints that dictate the amount of product we end a period with is equal to the amount we bought during that period, minus the amount we sold, plus the amount we previously had. $\mathbf{X} = \mathbf{B} - \mathbf{S} + \mathbf{X}_{prev}$. Where $\mathbf{X}_{prev}$ is a shift of $\mathbf{X}$

$$\mathbf{X}_{prev} = \begin{bmatrix} 500 & 500 & x_{11} & x_{12} & x_{13} \\ 500 & 500 & x_{21} & x_{22} & x_{23} \\ 500 & 500 & x_{31} & x_{32} & x_{33} \\ 500 & 500 & x_{41} & x_{42} & x_{43} \\ 500 & 500 & x_{51} & x_{52} & x_{53} \end{bmatrix} \tag{108}$$

The full optimization problem, with a rather large abuse of different notations and indices can be written as

$$\begin{aligned} \text{minimize} \quad & 150(\mathbf{1}^\top\mathbf{S}\mathbf{1}) - \textbf{trace}(\Lambda\mathbf{X}) - 5(\mathbf{1}^\top\mathbf{X}\mathbf{1} - 500 \times 5) && (109) \\ \text{subject to} \quad & \mathbf{X}_{o,t} \leq 1000 \quad \forall o \in O, t \in T && (110) \\ & \mathbf{B}_{1t} + \mathbf{B}_{2t} \leq 200 \quad \forall t \in T && (111) \\ & \mathbf{B}_{3t} + \mathbf{B}_{4t} + \mathbf{B}_{5t} \leq 250 \quad \forall t \in T && (112) \\ & \mathbf{S}^\top\mathbf{h} \preceq 6(\mathbf{S}^\top\mathbf{1}), \quad \mathbf{S}^\top\mathbf{h} \succeq 3(\mathbf{S}^\top\mathbf{1}) && (113) \\ & \mathbf{X} = \mathbf{B} - \mathbf{S} + \mathbf{X}_{prev} && (114) \\ & \mathbf{X}_{o,t}, \mathbf{B}_{o,t}, \mathbf{S}_{o,t} \geq 0 \quad \forall o \in O, t \in T && (115) \end{aligned}$$

3.3 Tuesday 02/04/2025

3.3.1 Standard Form

Standard form LPs follow the form

$$\begin{aligned} \text{minimize} \quad & c^\top x && (116)\\ \text{subject to} \quad & Ax = b && (117)\\ & x \geq 0 && (118) \end{aligned}$$

we can turn any LP into the above form with some tricks. An example problem, we want to turn the following optimization problem into standard form

$$\begin{aligned} \text{maximize} \quad & 3x_1 + 2x_2 - x_3 + x_4 && (119)\\ \text{subject to} \quad & x_1 + 2x_2 + x_3 - x_4 \leq 5 && (120)\\ & -2x_1 - 2x_4 - x_3 - x_4 \geq 1 && (121)\\ & x_1 \geq 0, x_2 \leq 0 && (122) \end{aligned}$$

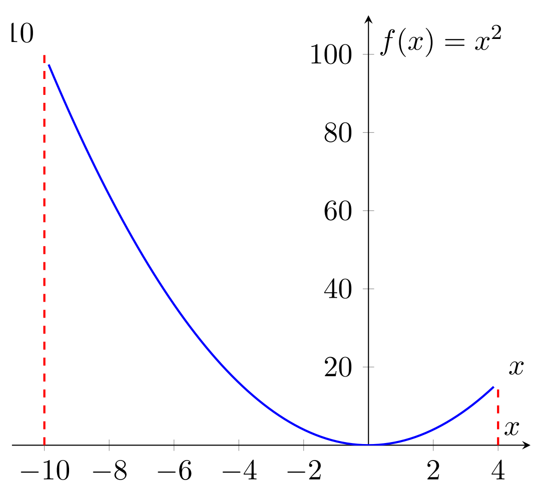
$$\begin{aligned} \text{minimize} \quad & -(3x_1 + 2(-x_2) - x_3 + x_4) && (123)\\ \text{subject to} \quad & x_1 + 2(-x_2) + x_3 - x_4 - s_1 = 5 && (124)\\ & 2x_1 - 2x_4 - x_3 - x_4 + s_2 = 1 && (125)\\ & x_1, x_2, s_1, s_2 \geq 0 && (126) \end{aligned}$$

In order to bound variables x_3, x_4, we can have them equal $x_3 = x_3^+ - x_3^-$.

3.3.2 Active Constraints

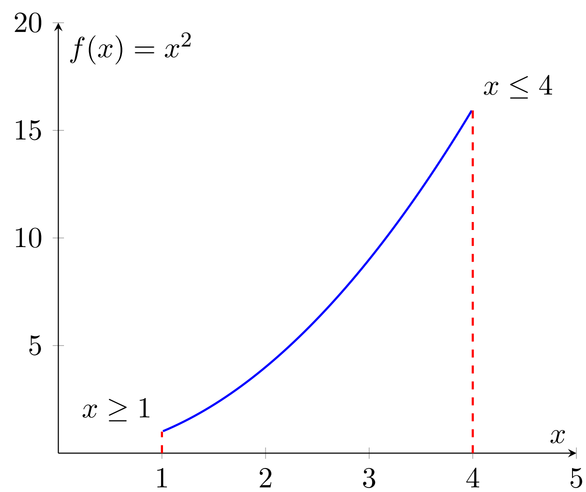
Consider the following optimization problem

$$\begin{aligned} \text{minimize} \quad & x^2 && (127)\\ \text{subject to} \quad & x \leq 4 && (128)\\ & x \geq -10 && (129) \end{aligned}$$



Both of the constraints in the above problem have no effect on the optimal solution. If they were removed, the answer to the problem would still be the same. These are inactive constraints. However, if we consider the next optimization problem

$$\text{minimize} \quad x^2 \tag{130}$$
$$\text{subject to} \quad x \leq 4 \tag{131}$$
$$x \geq 1 \tag{132}$$



The optimal solution is at $x = 1$. To formally define an active or inactive constraint, we can use the following:
Definition: Consider $g : \mathbb{R}^m \to \mathbb{R}^p$ and $h : \mathbb{R}^n \to \mathbb{R}^m$. For $i \in \{1, \ldots, p\}$inequality constraints,

$$g_i(x) \leq 0 \tag{133}$$

is said to be active in x^* if

$$g_i(x^*) = 0 \tag{134}$$

and inactive if

$$g_i(x^*) < 0 \tag{135}$$

The set of indices of active constraints at x^* is deonated as $\mathbb{A}(x^*)$

3.3.3 Feasible Direction

Using the same optimization problem as above in 3.3.2, we can illustrate the concept of a feasible direction. Starting at $x = -10$, the only feasible direction is to proceed in the positive x direction. For a general optimization problem and a feasible point $x \in \mathbb{R}^n$. A direction d is said to be feasible in x if there is $\eta > 0$ such that $x + \alpha d$ is feasible for any $0 \leq \alpha \leq \eta$. Our degree of freedom for movement is one less than the dimension of the total problem.
Theorem: Let $\mathbb{C}$ be a feasible convex set, and consider $x, y \in \mathbb{C}, y \neq x$. The direction $d = y - x$ is a feasible direction in $\mathbb{C}$ and $x + \alpha d$ is feasible for any $0 \leq \alpha \leq 1$.
A conclusion that can be drawn from the theorem above is that if we pick an interior point in $\mathbb{C}$ (not on the boundary of $\mathbb{C}$), any direction is feasible.

Theorem: Consider the optimization problem

$$\begin{aligned} \text{minimize} \quad & f(x) && (136) \\ \text{subject to} \quad & Ax = b && (137) \\ & x \geq 0 && (138) \end{aligned}$$

and let x^+ be a feasible point. A direction d is feasible iff $Ad = 0$ and $d_i \geq 0$ when $x_i^+ = 0$.

3.3.4 LP Region

The feasible region for a linear programming problem is a polyhedron. In canonical form, it can expressed as $\{x \in \mathbb{R}^n | Ax \leq 0\}$ and in standard form $\{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$.
Vertex: Let P be a polyhedron. A vector $x \in P$ is a vertex of P if it is impossible to find two vectors $y, z \in P, y \neq x, z \neq x$ such that x is a convex combination of the two points. A convex combination with $0 \leq \lambda \leq 1$ such that

$$x = \lambda y + (1 - \lambda) z \tag{139}$$

With this definition of a vertex, we can see that an optimal solution must be at a corner point.
Theorem: Let $P = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ be a polyhedron. $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Consider m linearly independent columns of A and the collection of those columns $B \in \mathbb{R}^{m \times m}$ and the remaining $n - m$ columns be $N \in \mathbb{R}^{m \times (n-m)}$.
In order to move around the columns of A, we have a permutation matrix P. This permutation matrix has a 1 in each column and each row. The permutation matrix will switch around the columns of matrix A.

$$AP = \begin{bmatrix} B \\ N \end{bmatrix} \tag{140}$$

$$x = P \begin{bmatrix} B^{-1} b \\ 0_{n-m} \end{bmatrix} \tag{141}$$

If $B^{-1} b \geq 0$, then this is a vertex

3.3.5 Basic solution

Let $P = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ be a polyhedron in standard form. $A \in \mathbb{R}^{m \times n}$, $n \geq m$, $x \in \mathbb{R}^n$, $Ax = b$. This vector x is called a basic solution if

- B is non singular (B is m linearly independent cols of A)
- $x_i = 0$ if $i \geq m$
- x $= B^{-1} b \geq 0$

3.4 Thursday 02/06/2025

3.4.1 Basic Feasible Soltuions

We define the following optimization problem

$$\begin{aligned} \text{minimize} \quad & c^T x && (142)\\ \text{subject to} \quad & Ax = 0 && (143)\\ & x \geq 0 && (144) \end{aligned}$$

We have a matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix} \tag{145}$$

We can get a basic solution by picking any two columns of the matrix and creating the basis with those two columns. We then solve for $x_B = B^{-1}b$ and set the other $x_N = 0$. A basic solution is the intersection of any two constraints. A basic feasible solution is the intersection of any two constraints that lie in the feasible region. By choosing x_1, x_2 in the basis from above, we get the following basic solution:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{146}$$

This solution above has three active constraints and one inactive constraint since $1 > 0$. This is a basic feasible solution since it obeys the constraints in the original problem. If we pick x_2, x_3 we get the following solution

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 0 \end{bmatrix} \tag{147}$$

This is a basic solution but is not a basic feasible solution since one of the entries is negative. Effectively, by finding the basic solutions, we are solving the $Ax = 0$ system and then at the end we check for feasibility against $x \geq 0$.

3.4.2 LP Solution Analysis

If all basic variables are > 0, then we have a unique solution to our problem. This is called non-degenerate. If any basic variables are $= 0$, then we have many values of x that will yield the optimal f^*. This is called degenerate.

The simplex algorithm works by traversing these basic solutions. It traverses along a basic direction. If x is a basic solution, then a basic direction is a feasible direction along the edges of the polyhedron. Consider the solution

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} \tag{148}$$

, we want to find a basic direction

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_B \\ \mathbf{d}_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_m \\ d_{m+1} \\ d_{m+2} \\ d_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_m \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{149}$$

. In order to find a basic direction, you set all non-basic variables $= 0$ except for 1. We want a feasible direction d where $Ad = 0$. We can partition A into two matrixes B, N.

$$\begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} \mathbf{d}_B \\ \mathbf{d}_N \end{bmatrix} = B\mathbf{d}_B + B\mathbf{d}_N = 0 \tag{150}$$

We can do some simplification on the middle section as such

$$= B\mathbf{d}_B + B\mathbf{d}_N = B\mathbf{d}_B + B \sum_{j=m+1}^{N} A_j d_j \tag{151}$$

$$= B\mathbf{d}_B + A_p \tag{152}$$

Where A_p is the p-th column of A.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} \alpha a & \beta b & \gamma c \\ \alpha d & \beta e & \gamma f \\ \alpha g & \beta h & \gamma i \end{bmatrix} = \alpha \begin{bmatrix} a \\ d \\ g \end{bmatrix} + \beta \begin{bmatrix} b \\ e \\ h \end{bmatrix} + \gamma \begin{bmatrix} c \\ f \\ i \end{bmatrix} \tag{153}$$

Therefore,

$$d_B = -B^{-1} A_p \tag{154}$$

$$\mathbf{d}_p = P \begin{bmatrix} d_{B_p} \\ d_{N_p} \end{bmatrix} \tag{155}$$

where $d_{B_p} = -B^{-1} A_p$ and d_{N_p} is a basis vector where the p-th entry is 1 and the rest are 0.

4 Chapter 4 - LP Solutions

4.1 Tuesday 02/11/2025

4.1.1 Simplex Algorithm

For an algorithm that solves an LP, we can define its input as an LP in the form

$$\begin{aligned} \text{minimize} \quad & c^\top x && (156) \\ \text{subject to} \quad & Ax = b && (157) \\ & x \succeq 0 && (158) \end{aligned}$$

where $c \in \mathbb{R}^n, b \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}$. The optimal solution lies at one of the vertices of the problem which we can represent as $J_1^0, \ldots, J_m^0$. The steps of the algoirthm include

1. $B = (A_{j1}, \ldots, A_{jm})$
2. Calculate the reduced cost in all basic directions as $\bar{c}_j = c_j - c_B^\top B^{-1} A_j, \quad \forall j \notin J^k$
3. If $\bar{c}_j \geq 0 \quad \forall j \notin J^k$ then this is the optimal solution. Otherwise, we have a collection of reduced costs for every non-basic variable we have. We then pick the calculate them one by one and pick the first one that is less than zero. We do this so we don't have to calculate each reduced cost to identify a descent direction.
4. Calculate the basic variables $x_B = B^{-1}b$.
5. Calculate the basic components of the basic direction $d_b = B^{-1}A_p$.
6. Move along $x_B + \alpha d_B$ where $\alpha_i = -\frac{x_{Bi}}{d_{bi}}$ if $d_{bi} < 0$ We will choose the total step α as the α_i that results in the smallest value.

For an example for part 3, consider a reduced cost vector:

$$\bar{c} = \begin{bmatrix} 2 \\ -2 \\ -4 \\ \vdots \\ -10000 \end{bmatrix}$$

In this case, we would pick the first value less than zero, which is -2.

4.1.2 Simplex Example

We can take an example problem

$$\begin{aligned} \text{minimize} \quad & -x_1 - 2x_2 && (159) \\ \text{subject to} \quad & x_1 + x_2 \leq 1 && (160) \\ & x_1 - x_2 \leq 1 && (161) \\ & x_1, x_2 \geq 0 && (162) \end{aligned}$$

which can be turned into standard form

$$\begin{aligned} \text{minimize} \quad & c^\top x && (163) \\ \text{subject to} \quad & Ax = b && (164) \\ & x \succeq 0 && (165) \end{aligned}$$

with parameters

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad c = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix} \tag{166}$$

$$J^0 = \{3, 4\} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad x_B = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = B^{-1}b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{167}$$

We solve for the reduced costs and get

$$\bar{c}_1 = c_1 - C_B^\top B^{-1} A_1 = -1 - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1 \tag{168}$$

$$\bar{c}_2 = c_2 - C_B^\top B^{-1} A_2 = -2 - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2 \tag{169}$$

$$d_{Np} = -B^{-1} A_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \tag{170}$$

$$d_p = \begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \end{bmatrix} \tag{171}$$

4.2 Thursday 02/13/2025

4.2.1 Geometric Interpretation of Steps

If we are at a vertex x, and we know a descent direction d, we want to figure out how far we should go along the direction to get a new $x_0 = x + \alpha d$. The direction may traverse multiple basic solutions, so we pick the lowest $\alpha_i = \frac{x}{d_i}$ so we ensure we do not violate the constraints.

4.2.2 Simplex Table

The simplex algorithm has some paramters calculated or held at each iteration

- Reduced cost: $\bar{c} = c^\top - c_B^\top B^{-1} A$
- Current iterate: $x_B = B^{-1}b$

- Basic direction: $d_B = -B^{-1}A_p$
- Objective function value: $-c^\top x$

Each of these must be stored. If we break it apart into a matrix we can store it with a matrix called the simplex table which is $\mathbb{R}^{(m+1)\times(n+1)}$
The simplex table is a convenient way to keep track of all the necessary information during the iterations of the simplex algorithm. It includes the coefficients of the constraints, the objective function, and the current solution.
For our example, the initial simplex table is:

	x_1	x_2	s_1	s_2	RHS
s_1	1	1	1	0	1
s_2	1	-1	0	1	1
Obj	-1	-2	0	0	0

Here, s_1 and s_2 are the slack variables added to convert the inequalities into equalities. The "RHS" column represents the right-hand side of the constraints.
The simplex algorithm has two parts, phase 1 is finding a basic feasible solution using the simplex algorithm and phase 2 is solving given the basic feasible solution.

4.2.3 Updating the Simplex table

In order to update a simplex table that has a feasible solution, we can use the example below
For a simplex table with 3 constraints and 6 variables, we can set it up as follows:

	x_1	x_2	x_3	x_4	x_5	x_6	RHS
x_1	0	1.5	1	1	-0.5	0	10
x_4	1	0.5	1	0	0.5	0	10
x_6	0	1	-1	0	-1	1	0
Obj	0	-7	-2	0	5	0	100

Here, we pick x_2 as the entering variable since it is the first index with a negative value. We calculate α for each basic variable to calculate which is the one leaving the simplex table.

- $\alpha_1 = \frac{20}{25}$
- $\alpha_2 = \frac{20}{25}$
- $\alpha_1 = \frac{1}{1}$

We pivot the table on row x_6, and column x_2.