Optimization

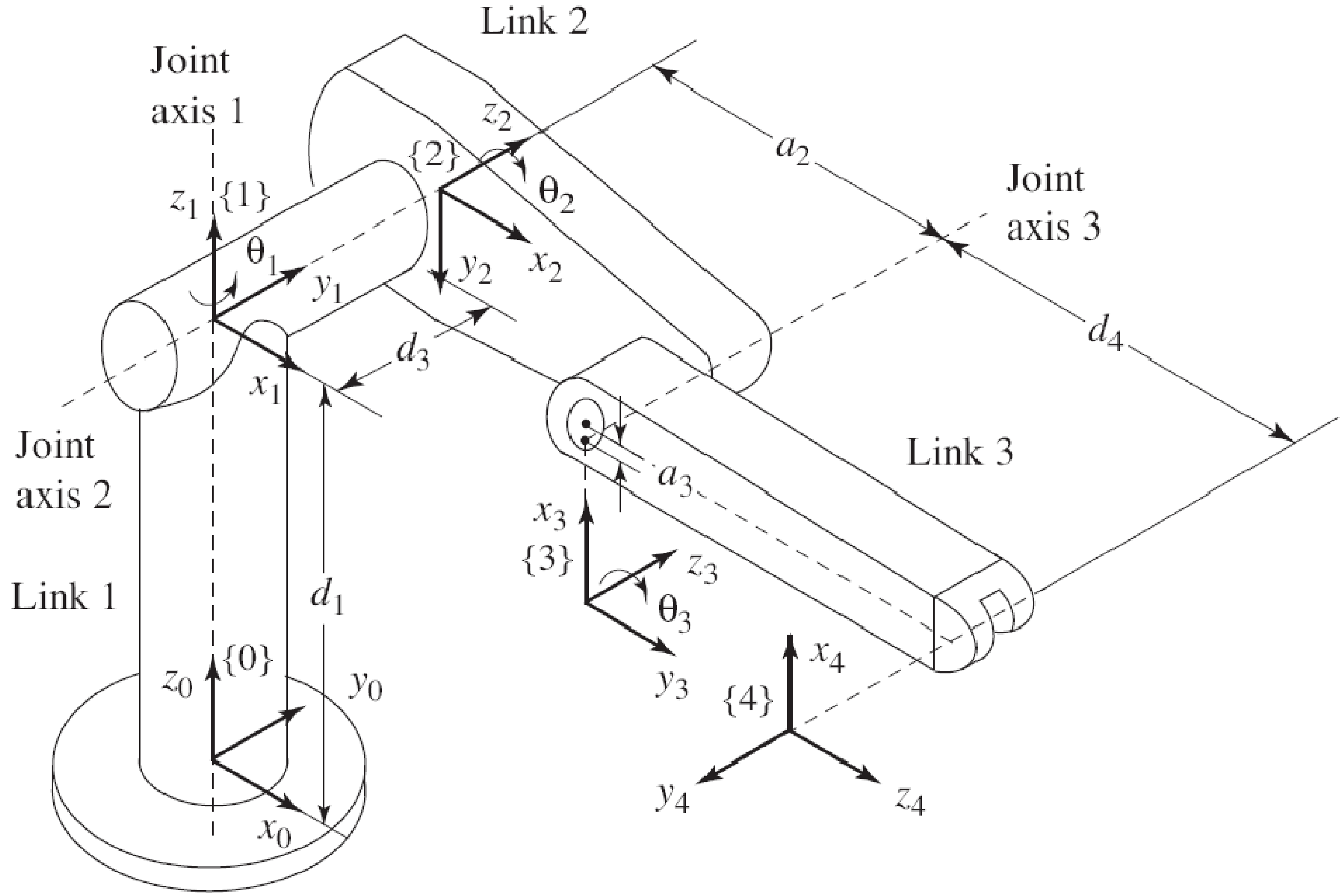
In an *optimization problem*, the objective is to optimize (maximize or minimize) some function f. This function f is called the objective function.

For example, an objective function f to be maximized may be the revenue in a production of TV sets, the yield per minute in a chemical process, the hourly number of customers served in some office, the hardness of steel, or the tensile strength of a rope.

Similarly, we may want to minimize f if f is the cost per unit of producing certain cameras, the operating cost of some power plant, the daily loss of heat in a heating system, the idling time of some lathe, or the time needed to produce a fender.

Unconstrained Optimization: An Example

Inverse kinematics of a robotic manipulator: Given tip position $\{p_x, p_y, p_z\}$, find joint rotations $\{\theta_1, \theta_2, \theta_3\}$.



Courtesy of Professor Wu-Sheng Lu of University of Victoria, Canada

Unconstrained Optimization: An Example (cont.)

A solution via forward kinematics

$$
\begin{aligned}
c_1\left(a_2c_2 + a_3c_{23} - d_4s_{23}\right) - d_3s_1 &= p_x \\
s_1\left(a_2c_2 + a_3c_{23} - d_4s_{23}\right) + d_3c_1 &= p_y \\
d_1 - a_2s_2 - a_3s_{23} - d_4c_{23} &= p_z
\end{aligned}
$$

$$\Updownarrow$$

$$
\begin{aligned}
f_1(\Theta) &\triangleq c_1\left(a_2c_2 + a_3c_{23} - d_4s_{23}\right) - d_3s_1 - p_x = 0 \\
f_2(\Theta) &\triangleq s_1\left(a_2c_2 + a_3c_{23} - d_4s_{23}\right) + d_3c_1 - p_y = 0 \\
f_3(\Theta) &\triangleq d_1 - a_2s_2 - a_3s_{23} - d_4c_{23} - p_z = 0
\end{aligned}
$$

$$\Theta = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 \end{bmatrix}^T$$

Courtesy of Professor Wu-Sheng Lu of University of Victoria, Canada

Unconstrained Optimization: An Example (cont.)

An optimization-based approach

$$\left\{ \begin{array}{l} f_1(\Theta) = 0 \\ f_2(\Theta) = 0 \\ f_3(\Theta) = 0 \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} f_1^2(\Theta) = 0 \\ f_2^2(\Theta) = 0 \\ f_3^2(\Theta) = 0 \end{array} \right\} \Leftrightarrow \sum_{i=1}^{3} f_i^2(\Theta) = 0$$

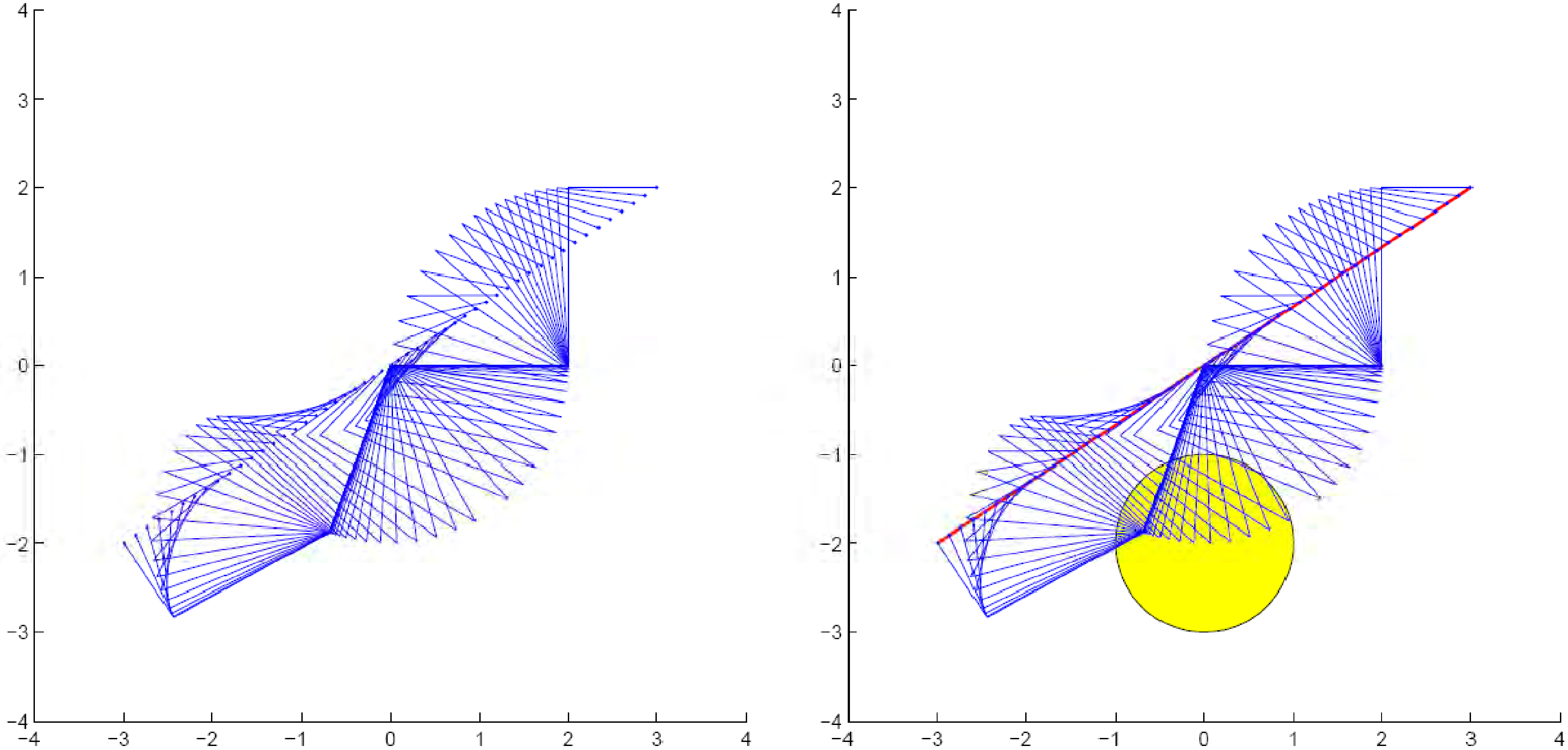
$$\Rightarrow \qquad \underset{\Theta}{\text{minimize}} \;\; F(\Theta) = \sum_{i=1}^{3} f_i^2(\Theta)$$

Advantages of the approach:

- it works regardless of the relation of the number of equations versus the number of unknowns.
- it offers a "best" approximate solution if no exact solutions exist.

Courtesy of Professor Wu-Sheng Lu of University of Victoria, Canada

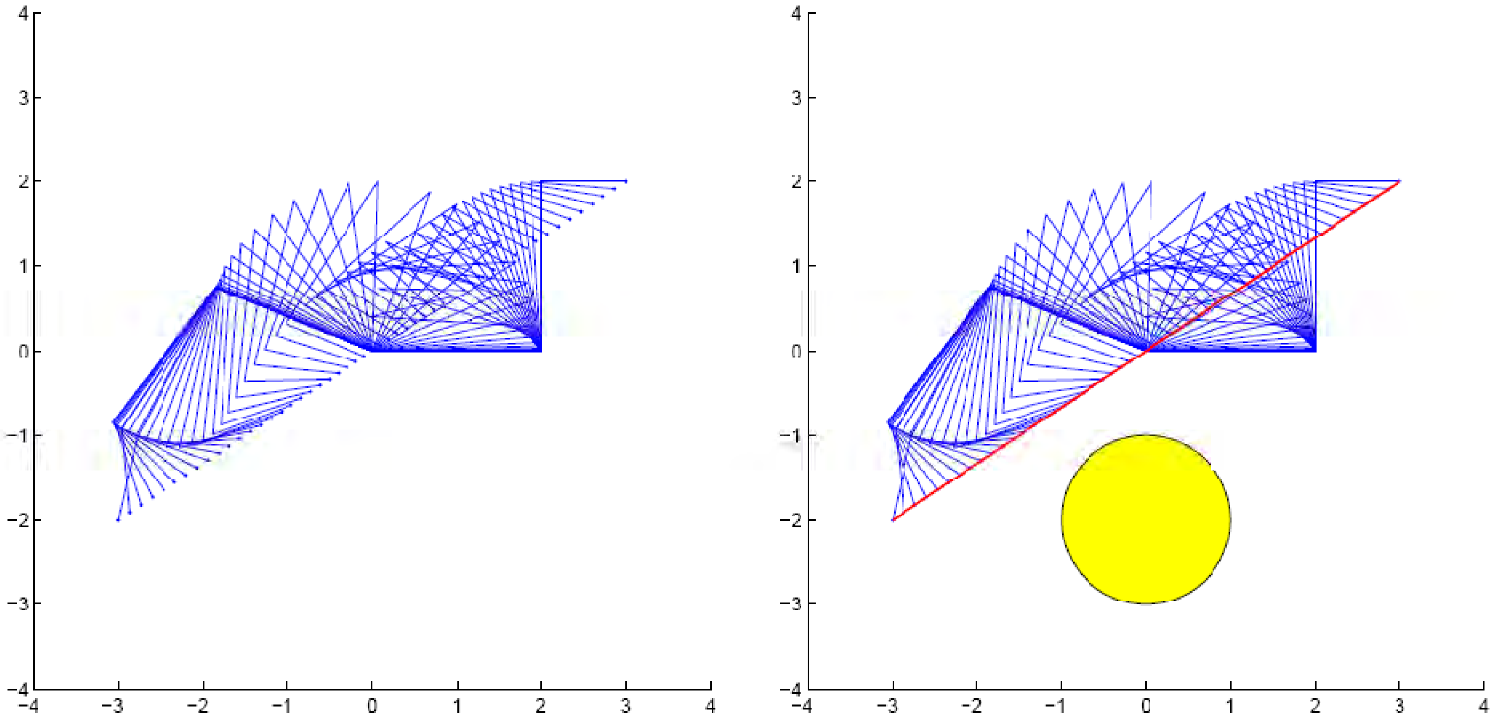
Unconstrained Optimization: An Example (cont.)



Constrained Optimization: An Example

- A constrained-optimization based path planning for obstacle avoidance

$$\text{minimize} \qquad F = \int_{t_0}^{t_1} g(\theta, t)\, dt$$
$$\text{subject to: } X(t) = f(\theta(t)) \quad \text{(kinematics)}$$



Courtesy of Professor Wu-Sheng Lu of University of Victoria, Canada

Background

In most optimization problems the objective function f depends on several variables

$$x_1, \cdots, x_n.$$

There are called *control variables* because we can "control" them, e.g.,

1) The yield of a chemical process may depend on pressure x_1 and temperature x_2.

2) The efficiency of a certain air conditioning system may depend on temperature x_1, air pressure x_2, moisture content x_3, cross-sectional area of outlet x_4, etc.

Background

Optimization theory develops methods for optimal choices of $x_1, x_2, \ldots, x_n$, which maximize (or minimize) the objective function f, that is methods for finding optimal values of $x_1, x_2, \ldots, x_n$.

In many problems the choice of values of $x_1, x_2, \ldots, x_n$ is not entirely free but is subject to some constraints, that is, additional conditions arising from the nature of the problem and the variables.

For example, if x_1 is production cost, then $x_1 \geq 0$ and there are many other variables (time, weight, distance travelled by a salesman, etc.) that can take nonnegative values only. Constraints can also have the form of equations (instead of inequalities).

Unconstrained Optimization

Let us first consider unconstrained optimization in the case of a real-valued function $f(x_1, x_2, \ldots, x_n)$. We also write $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ and $f(\mathbf{x})$, for convenience.

f has a minimum at point $\mathbf{x} = \mathbf{X}_0$ in a region $\boldsymbol{R}$ if $f(\mathbf{x}) \geq f(\mathbf{X}_0)$ for $\mathbf{x}$ in $\boldsymbol{R}$.

Similarly, f has a maximum at $\mathbf{X}_0$ if $f(\mathbf{x}) \leq f(\mathbf{X}_0)$ for all $\mathbf{x}$ in $\boldsymbol{R}$. Minima and maxima are called *extrema*.

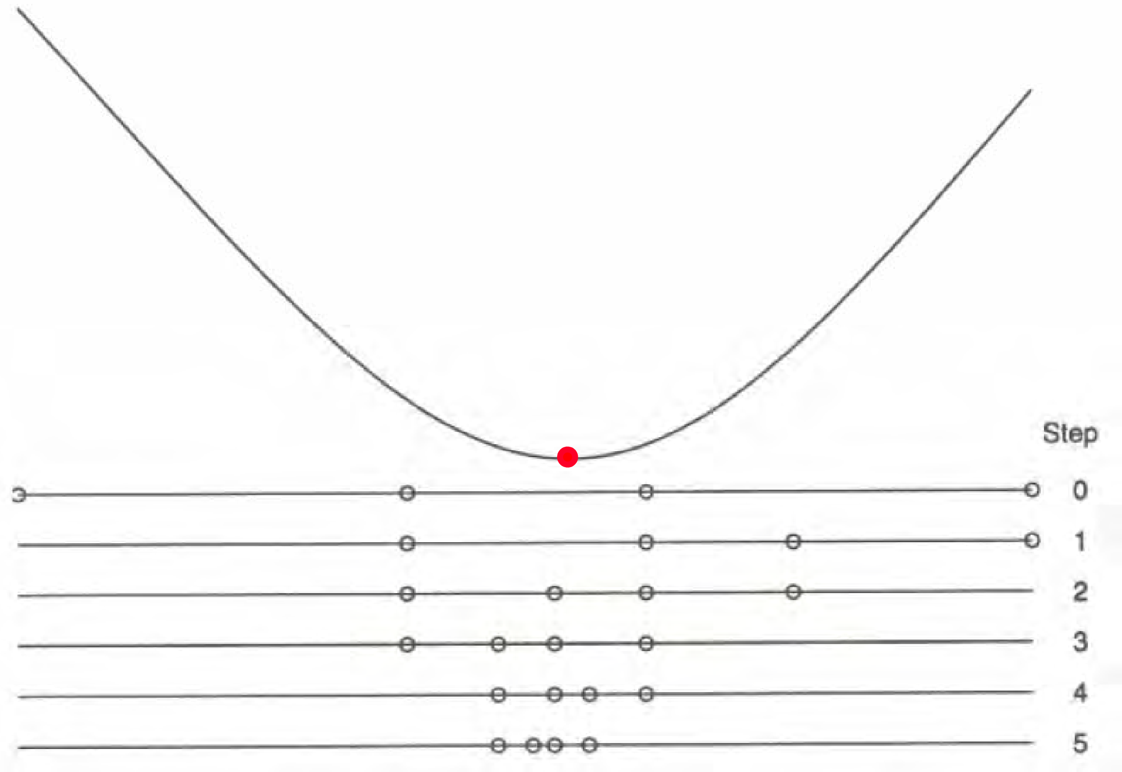
f is said to have a *local minimum* at $\mathbf{X}_0$ if $f(\mathbf{x}) \geq f(\mathbf{X}_0)$ for all $\mathbf{x}$ in a neighbourhood of $\mathbf{X}_0$, for all $\mathbf{x}$ satisfying

$$|\mathbf{x} - \mathbf{X}_0| < r$$

where $r > 0$ is sufficiently small.

Example:

Minimizing a Function of a Single Variable



Golden Section Search

Unconstrained Optimization

Question: How can one reach the valley bottom (minima) fastest?

Answer: Take the path in the most downhill direction. Mathematically, this is the direction of the negative gradient.

Unconstrained Optimization (cont.)

If f is differentiable and has an extremum at a point $\mathbf{X}_0$, then the partial derivatives $\partial f/\partial x_1,\cdots,\partial f/\partial x_n$ must be zero at $\mathbf{X}_0$. Thus its gradient

$$\nabla f(\mathbf{X}_0)=\begin{pmatrix}\partial f/\partial x_1\\ \vdots\\ \partial f/\partial x_n\end{pmatrix}_{\mathbf{x}=\mathbf{X}_0}=\mathbf{0}. \qquad (1)$$

A point $\mathbf{X}_0$ at which $\nabla f(\mathbf{X}_0)=\mathbf{0}$ is called a *stationary point* (valley bottom) of f.

Condition (1) is necessary for an extremum of f at $\mathbf{X}_0$ in the interior of $\boldsymbol{R}$, but is not sufficient. In practice, solving (1) will often be difficult. For this reason, one generally prefers solution by iteration, by search processes that start at some point and move stepwise to points at which f is smaller (if a minimum of f is wanted) or bigger (in the case of maximum).

Cauchy's Method

Cauchy's *method of steepest descent* or *gradient method* is one such popular method. However, convergence can sometimes be slow. Given a multivariable function $f(x)$, we examine its Taylor expansion (obtained from Taylor series)

$$f(\mathbf{x}+t\cdot\delta\mathbf{x}) = f(\mathbf{x}) + t\nabla f(\mathbf{x})^{\mathrm{T}}\delta\mathbf{x} + \cdots,$$

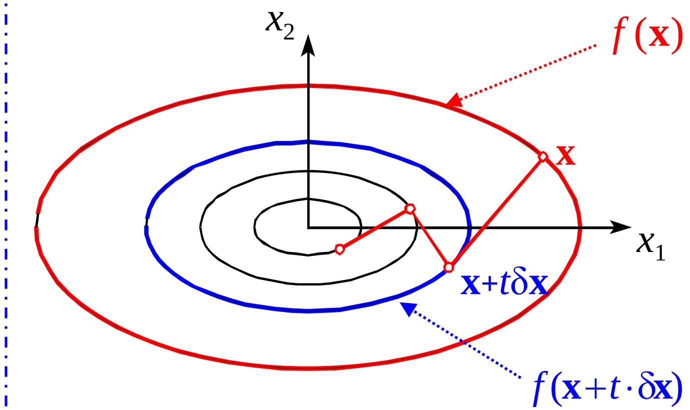
$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \delta\mathbf{x} = \begin{pmatrix} \delta x_1 \\ \vdots \\ \delta x_n \end{pmatrix}, \nabla f(\mathbf{x}) = \begin{pmatrix} \partial f/\partial x_1 \\ \vdots \\ \partial f/\partial x_n \end{pmatrix}$$

Thus, we can expect a decrease in value of f if we set the step direction $\delta\mathbf{x}$ to be

$$\delta\mathbf{x} = -\nabla f(\mathbf{x})$$

and the step size $t \geq 0$. This is because

$$f(\mathbf{x}-t\nabla f(\mathbf{x})) \approx f(\mathbf{x}) - t\left[\underbrace{\nabla f(\mathbf{x})^{\mathrm{T}}\nabla f(\mathbf{x})}_{\geq 0}\right] \leq f(\mathbf{x})$$



Cauchy's Method (cont.)

Why $\nabla f(\mathbf{x})^{\mathrm{T}} \nabla f(\mathbf{x}) \geq 0$? Noting that

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix} \Rightarrow \nabla f(\mathbf{x})^{\mathrm{T}} = \begin{pmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{pmatrix}$$

and

$$\nabla f(\mathbf{x})^{\mathrm{T}} \nabla f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{pmatrix} \begin{pmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix} = \left(\frac{\partial f}{\partial x_1} \right)^2 + \cdots + \left(\frac{\partial f}{\partial x_n} \right)^2 \geq 0$$

thus, we have

$$f(\mathbf{x} - t\nabla f(\mathbf{x})) \approx f(\mathbf{x}) - t\left[\nabla f(\mathbf{x})^{\mathrm{T}} \nabla f(\mathbf{x}) \right] \leq f(\mathbf{x}) \qquad (2)$$

Basic Idea of Cauchy's Method

Now suppose that we start from an initial point, say $\mathbf{x}_0$, and an appropriate step size $t_1 \ge 0$ and let $\mathbf{x}_1 = \mathbf{x}_0 - t_1 \nabla f(\mathbf{x}_0)$. From (2), we have

$$f(\mathbf{x}_1) = f(\mathbf{x}_0 - t_1 \nabla f(\mathbf{x}_0)) \approx f(\mathbf{x}_0) - t_1 \left[\nabla f(\mathbf{x}_0)^{\mathrm{T}} \nabla f(\mathbf{x}_0) \right] \le f(\mathbf{x}_0)$$

Let us move on from $\mathbf{x}_1$ by defining $\mathbf{x}_2 = \mathbf{x}_1 - t_2 \nabla f(\mathbf{x}_1)$ with an appropriately chosen step size t_2. Again from (2), we have

$$f(\mathbf{x}_2) = f(\mathbf{x}_1 - t_2 \nabla f(\mathbf{x}_1)) \approx f(\mathbf{x}_1) - t_2 \left[\nabla f(\mathbf{x}_1)^{\mathrm{T}} \nabla f(\mathbf{x}_1) \right] \le f(\mathbf{x}_1) \le f(\mathbf{x}_0)$$

By repeating this process, we obtain an iterative sequence or scheme:

$$\mathbf{x}_i = \mathbf{x}_{i-1} - t_i \nabla f(\mathbf{x}_{i-1}) \qquad (3)$$

such that the resulting sequence has the following property:

$$f(\mathbf{x}_0) \ge f(\mathbf{x}_1) \ge \cdots \ge f(\mathbf{x}_{i-1}) \ge f(\mathbf{x}_i) \ge \cdots \ge f(\mathbf{X}_0)$$

The sequence generated by (3) would converge to $\mathbf{X}_0$ as i tends to ∞.

Cauchy's Method – Selection of Step Size

The issue on how to select an appropriate step size is quite complicated. In principle, we can fix the step size t to be some appropriate small positive scalar, say t^*, and carry on the iteration:

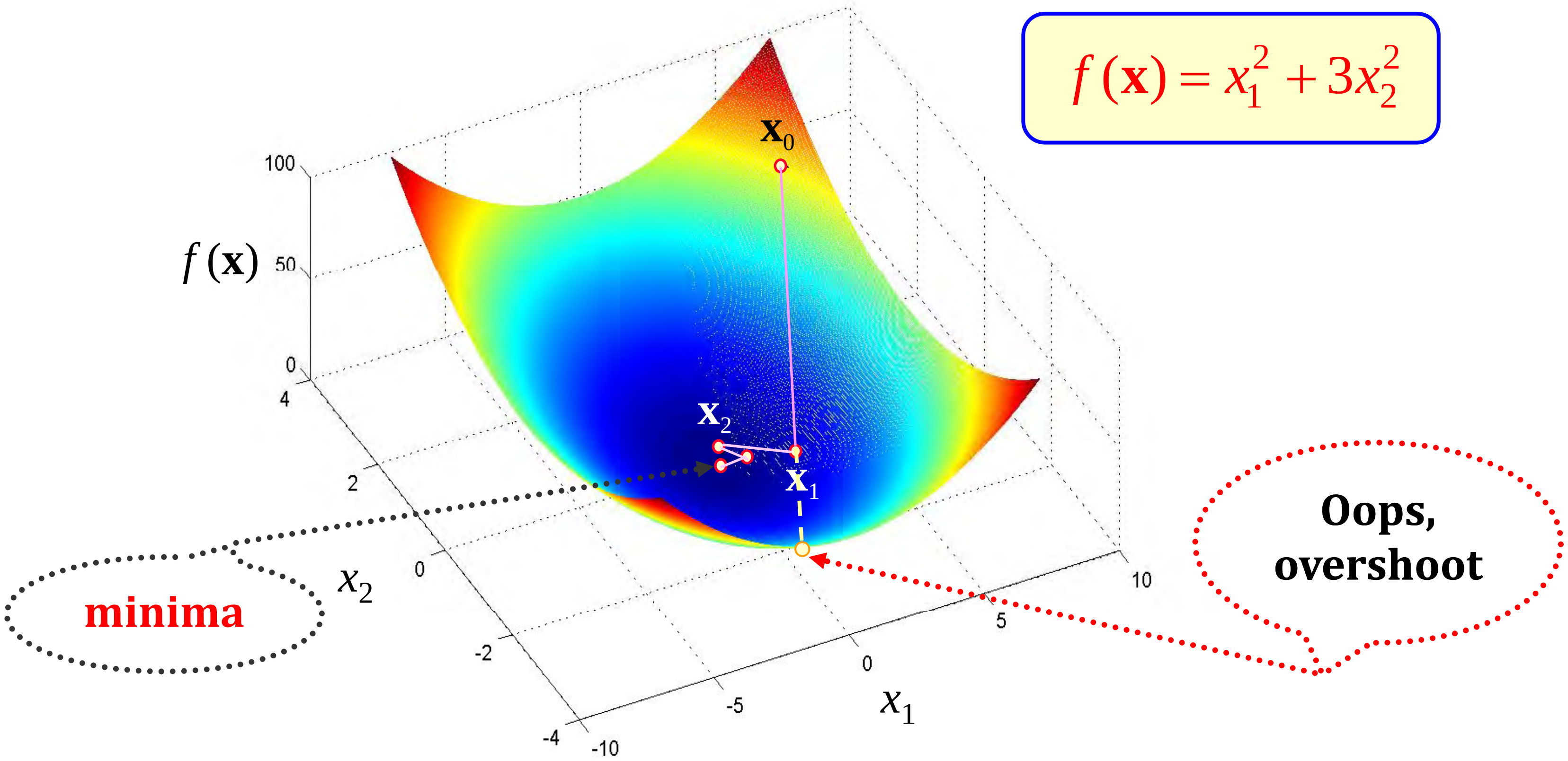
$$\mathbf{x}_i = \mathbf{x}_{i-1} - t^* \nabla f(\mathbf{x}_{i-1}) \tag{4}$$

Such an iteration might not be efficient, but surely works. Alternatively, we can also determine the step size t_i and the corresponding point

$$\mathbf{x}_i = \mathbf{x}_{i-1} - t_i \nabla f(\mathbf{x}_{i-1})$$

at which the function $f(\mathbf{x}_i)$ is the smallest (a bit closer to the minima ☺) among all the possible choices of step size t. We illustrate this in an example…

An Illustrative Example



Example

Determine a minimum of $f(\mathbf{x}) = x_1^2 + 3x_2^2$, starting from $\mathbf{x}_0 = (6 \quad 3)'$ using the method of steepest descent. Clearly, inspection shows that $f(x)$ has a minimum at $\mathbf{0}$. Knowing the solution gives us a better feeling of how the method works. We first obtain the gradient

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 6x_2 \end{pmatrix}$$

and the iteration scheme

$$\mathbf{x}_i(t) = \mathbf{x}_{i-1} - t\nabla f(\mathbf{x}_{i-1}) = \begin{pmatrix} x_{1,i-1} \\ x_{2,i-1} \end{pmatrix} - t \begin{pmatrix} 2x_{1,i-1} \\ 6x_{2,i-1} \end{pmatrix} = \begin{pmatrix} (1-2t)x_{1,i-1} \\ (1-6t)x_{2,i-1} \end{pmatrix} \qquad (5)$$

Note that $\mathbf{x}_i$ is depended on the choice of the step size t, and $x_{1,i-1}$ and $x_{2,i-1}$ are the values of x_1 and x_2 corresponding to $\mathbf{x}_{i-1}$. We are now to select a step size t_i such that the resulting $f(\mathbf{x}_i)$ is the smallest among all the choices of t.

Optimal Selection of Step Size

$$f(\mathbf{x}) = x_1^2 + 3x_2^2$$

From (5), we have

$$f(\mathbf{x}_i(t)) = f\begin{pmatrix}(1-2t)x_{1,i-1}\\(1-6t)x_{2,i-1}\end{pmatrix} = (1-2t)^2 x_{1,i-1}^2 + 3(1-6t)^2 x_{2,i-1}^2$$

To find the small value of f above with respect to t, we take

$$\frac{d}{dt} f(\mathbf{x}_i(t)) = \frac{d}{dt}\left[(1-2t)^2 x_{1,i-1}^2\right] + \frac{d}{dt}\left[3(1-6t)^2 x_{2,i-1}^2\right]$$

$$= 2\cdot(1-2t)x_{1,i-1}^2\cdot(-2) + 2\cdot 3(1-6t)x_{2,i-1}^2\cdot(-6)$$

$$= -4(1-2t)x_{1,i-1}^2 - 36(1-6t)x_{2,i-1}^2$$

$$= \left(8x_{1,i-1}^2 + 216x_{2,i-1}^2\right)t - \left(4x_{1,i-1}^2 + 36x_{2,i-1}^2\right) = 0$$

$$\Rightarrow \quad t_i = \frac{x_{1,i-1}^2 + 9x_{2,i-1}^2}{2x_{1,i-1}^2 + 54x_{2,i-1}^2} \quad \Leftarrow$$

the optimal choice of step size

Starting from $\mathbf{x}_0 = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$, we compute the values for x_1 and x_2 as listed in the table below and plotted in the figures on the next page.

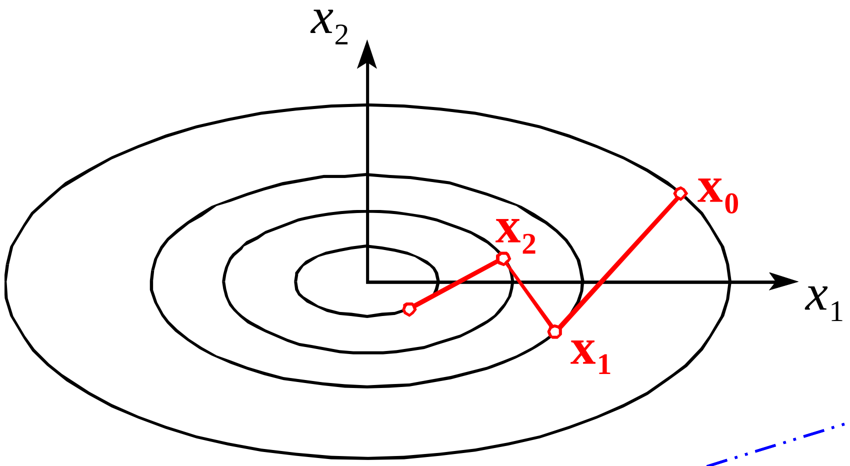
Step i	$\mathbf{x}_i$: $x_{1,i}$	$\mathbf{x}_i$: $x_{2,i}$	**Step Size** t_{i+1}
0	6.000	3.000	0.210
1	3.484	−0.774	0.310
2	1.327	0.664	0.210
3	0.771	−0.171	0.310
4	0.294	0.147	0.210
5	0.170	−0.038	0.310
6	0.065	0.032	**fast…**

$$t_i = \frac{x_{1,i-1}^2 + 9x_{2,i-1}^2}{2x_{1,i-1}^2 + 54x_{2,i-1}^2}$$

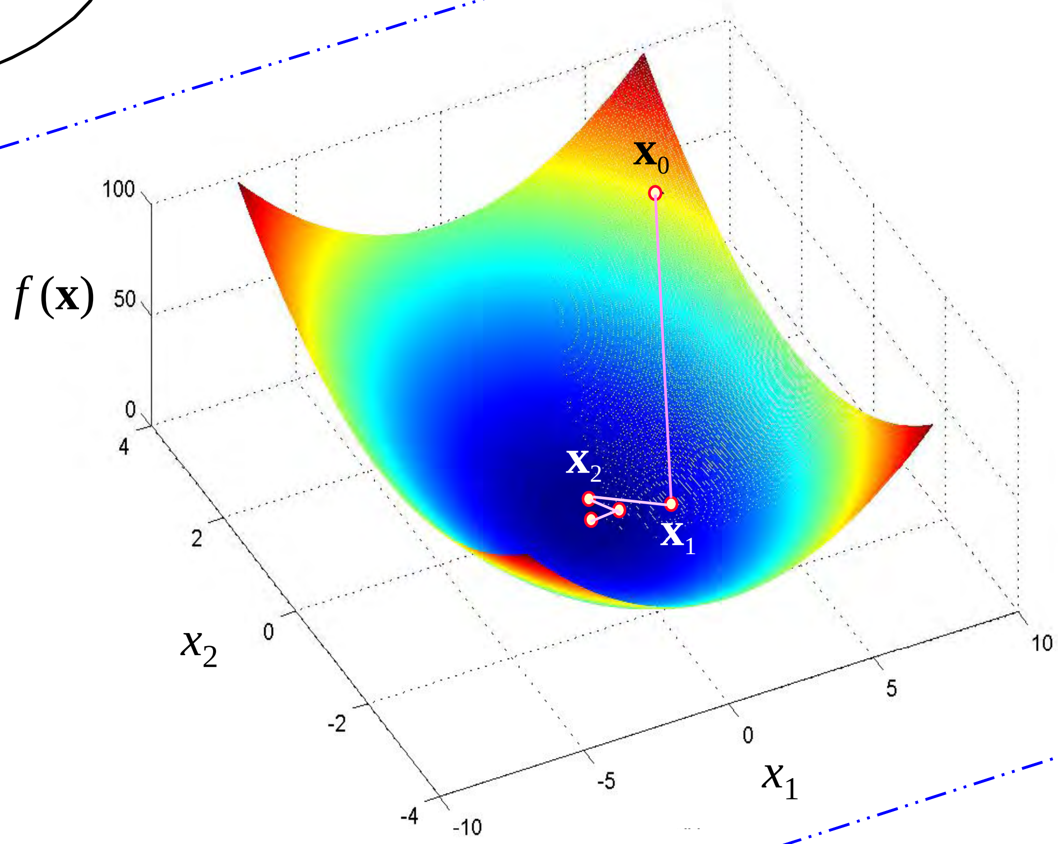
$$\nabla f(\mathbf{x}_{i-1}) = \begin{pmatrix} 2x_{1,i-1} \\ 6x_{2,i-1} \end{pmatrix}$$

$$\mathbf{x}_i = \mathbf{x}_{i-1} - t_i \nabla f(\mathbf{x}_{i-1})$$

steepest1.m



2D View



3D View

Example (cont.) – Solving it with a fixed step size

Step i	$\mathbf{x}_i$: $x_{1,i}$	$\mathbf{x}_i$: $x_{2,i}$	Fixed Step Size t^*
0	6.0000	3.0000	0.15
1	4.2000	0.3000	0.15
2	2.9400	0.0300	0.15
3	2.0580	0.0030	0.15
4	1.4406	0.0003	0.15
⋮	⋮	⋮	0.15
20	0.0048	3×10^{-20}	**slow…**

$$\nabla f(\mathbf{x}_{i-1}) = \begin{pmatrix} 2x_{1,i-1} \\ 6x_{2,i-1} \end{pmatrix}$$

$$t^* = 0.15$$

$$\mathbf{x}_i = \mathbf{x}_{i-1} - t^* \nabla f(\mathbf{x}_{i-1})$$

The drawback of the method with a fixed step size is: it is generally slow in converging to the solution.

steepest2.m

Summary of Unconstrained Optimization (Cauchy's Method)

Given a real valued nonlinear function $f(\mathbf{x}) = f(x_1, \cdots, x_n)$ and an initial point $\mathbf{x}_0$, the problem is to find a solution $\mathbf{X}_0$ such that $f(\mathbf{X}_0)$ is optimal (either minimum or maximum)

↓

Compute the gradient of $f(\mathbf{x})$, i.e., $\nabla f(\mathbf{x})$

↓

Design and perform an iterative scheme: $\mathbf{x}_i = \mathbf{x}_{i-1} - t_i \nabla f(\mathbf{x}_{i-1})$

↓

If the iterative scheme converges and stops at a certain step, say k, then $\mathbf{X}_0 \approx \mathbf{x}_k$ and the minimum value of $f(\mathbf{x})$ is then approximately given by $f(\mathbf{x}_k)$, i.e., $f(\mathbf{X}_0) \approx f(\mathbf{x}_k)$

Linear Programming

Linear programming (or linear optimization) consists of methods for solving optimization problems *with constraints* in which the objective function f is a *linear* function of the control variables $x_1, x_2, \ldots, x_n$.

Problems of this type arise in production, distribution of goods economics, and approximation theory.

Example

A Container Production Optimization Problem:

Suppose that in producing two types of containers K and L one uses two machines M_1 and M_2. To produce a container K, one needs M_1 two minutes and M_2 four minutes. Similarly, L occupies M_1 eight minutes and M_2 four minutes. The net profit for a container K is \$29 and for L it is \$45.

Determine the production plan that maximizes the net profit.

Example (cont.)

Problem Formulation:

If we produce x_1 containers K and x_2 containers L per hour, the profit per hour is

$$f(x_1, x_2) = 29\, x_1 + 45\, x_2.$$

The constraints are

$$\begin{aligned} 2x_1 + 8x_2 &\le 60 \qquad \text{(resulting from machine } M_1\text{)} \\ 4x_1 + 4x_2 &\le 60 \qquad \text{(resulting from machine } M_2\text{)} \\ x_1 &\ge 0 \\ x_2 &\ge 0 \end{aligned}$$

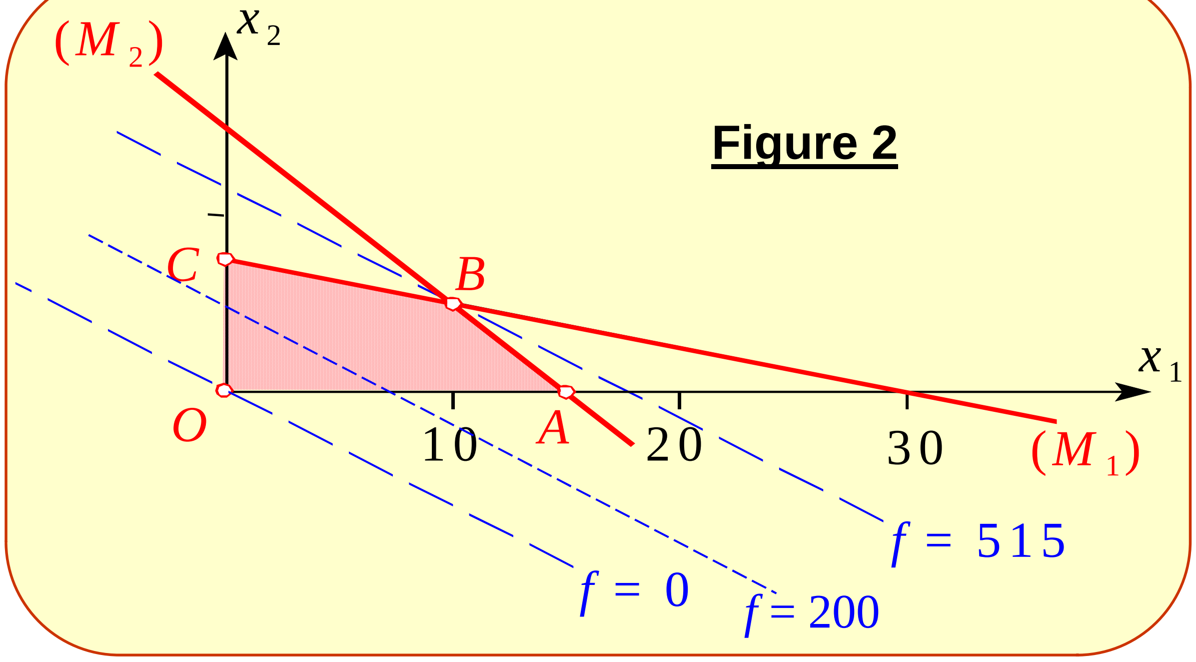
Figure 2 shows these constraints (x_1, x_2) must lie in the first quadrant and below or on the straight line $2\,x_1 + 8\,x_2 = 60$ as well as below or on the line $4\,x_1 + 4\,x_2 = 60$. Thus, (x_1, x_2) is restricted to the quadrangle $OABC$.

We have to find (x_1, x_2) in $OABC$ such that $f\,(x_1, x_2)$ is maximum.

Now $f\,(x_1,x_2) = 0$ gives $x_2 = -\,(29/45)\,x_1$ (see Fig. 2). The lines $f\,(x_1, x_2) = \text{constant}$ are parallel to that line. We see that B, that is, $x_1 = 10$ and $x_2 = 5$, gives the optimum $f\,(10, 5) = 515$.

Hence the answer is that the optimal production plan that maximizes the profit is achieved by producing containers K and L in the ratio 2:1, the maximum profit being **$515** per hour.

Figure 2



Simplex Method

In practice, linear programming problems might contain many more variables than the two variables considered in the previous example. A computational method is then required to solve the problem. One such technique is the **Simplex Method**. We illustrate such a method through examples…

Example

Maximize $f = x_1 + 4x_2$ subject to constraints

$$-x_1 + 2x_2 \le 6$$
$$5x_1 + 4x_2 \le 40$$
$$x_1, x_2 \ge 0$$

Idea: Introduce slack variables w_1, w_2. The constraints then become

$$-x_1 + 2x_2 + w_1 = 6$$
$$5x_1 + 4x_2 + w_2 = 40$$
$$x_1, x_2, w_1, w_2 \ge 0$$

and the objective function $f - x_1 - 4x_2 = 0$

Solution: Simplex table (to be constructed on the next page)…

Formulation of Simplex Method

The problem is equivalent to finding one of the solutions of the following set of linear equations:

$$\begin{aligned} 0\cdot f - x_1 + 2x_2 + w_1 + 0\cdot w_2 &= 6 \\ 0\cdot f + 5x_1 + 4x_2 + 0\cdot w_1 + w_2 &= 40 \\ f - x_1 - 4x_2 + 0\cdot w_1 + 0\cdot w_2 &= 0 \end{aligned}$$

subject to the constraints, $x_1, x_2, w_1, w_2 \ge 0$, such that f is maximum. From what we have learnt in linear algebra, we know solutions to the above linear equations remain unchanged with the following 3 basic operations (**B.O.**):

B.O.1: Interchange of two equations

B.O.2: Multiplication of an equation by a nonzero constant

B.O.3: Addition of a multiple of one equation to another equation

Key Idea of Simplex Method

The key idea of Simplex Method is to carry out a series of these basic operations (mainly B.O.2 and B.O.3) on the equations to transfer them into an certain form for which the desired solution (the solution corresponding to the maximum f can be easily observed and deduced).

We will illustrate this idea through specific examples…

Simplex Table:

Similar to the Matrix Form of Linear Equations

Basis	x_1	x_2	w_1	w_2	b	check
w_1	–1	2	1	0	6	8
w_2	5	4	0	1	40	50
f	–1	–4	0	0	0	–5

Problem variables: x_1, x_2; Slack variables: w_1, w_2; Constants: b

unity matrix

$$0 \cdot f - x_1 + 2x_2 + w_1 + 0 \cdot w_2 = 6$$
$$0 \cdot f + 5x_1 + 4x_2 + 0 \cdot w_1 + w_2 = 40$$
$$f - x_1 - 4x_2 + 0 \cdot w_1 + 0 \cdot w_2 = 0$$

Simplex Method (cont.)

The check column on the right hand side is included to provide a check on the numerical calculations as we develop the simplex. For each row, total up the entries in that row and enter the sum in the check column.

We are to perform a series of basic operations (B.O.2 and B.O.3) such that the coefficients associated with the objective function are all nonnegative. For the given example, it will be seen soon that the objective function can be transformed into:

$$f + 0 \cdot x_1 + 0 \cdot x_2 + \frac{8}{7} w_1 + \frac{3}{7} w_2 = 24 \quad \Rightarrow \quad f = 24 - \frac{8}{7} w_1 - \frac{3}{7} w_2, \quad w_1, w_2 \ge 0$$

Clearly, the maximum of f is **24** (by setting $w_1 = w_2 = 0$).

Basis	x_1	x_2	w_1	w_2	b	check
w_1	–1	2	1	0	6	8
w_2	5	4	0	1	40	50
f	–1	–4	0	0	0	–5

key row

key column

Steps:

1. *Key column*: Select the most negative entry in the index row; in this case –4.
2. *Key row*: Divide the entry in the b-column by the positive entry in the key column. The smallest positive ratio determines the key row.
3. The entry at the intersection of the key column and the key row is called the *pivot*.
4. Divide each entry in the key row by the pivot to reduce the pivot to a unit pivot, which we then circle. The revised key row is now called the *main row*.

B.O.2.

Basis	x_1	x_2	w_1	w_2	b	check	
w_1	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	3	4	Main row
w_2	5	4	0	1	40	50	
f	–1	–4	0	0	0	–5	Index row

Key column

5. Use the main row to operate on the remaining rows to reduce the other entries in the key column to zero. The new entries can be calculated as follows:

New number = Old number – product of corresponding entries in key row and key column

For example, in the second row (w_2):

5 is replaced by $5-(-\frac{1}{2})(4) = 5+2 = 7$ B.O.3.

In the third row (f):

-1 is replaced by $-1-(-\frac{1}{2})(-4) = -1-2 = -3.$ B.O.3.

6. Confirm that the new values in the check column are indeed the sums of the entries in the corresponding rows. If not, there is a mistake somewhere in the calculation, which should be corrected.

7. Change of basic variables: Change the variable in the key column (x_2) with the basic variable in the main row (w_1).

Basis	x_1	x_2	w_1	w_2	b	check
x_2 ~~w_1~~	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	3	4
w_2	7	0	–2	1	28	34
f	–3	0	2	0	12	11

The basic variables are now x_2 and w_2, and the basic solution is thus $x_2 = 3$, $w_2 = 28$. However, the index row (f) still contains a negative entry, and therefore this is not the optimum solution.

8. Repeat steps 1 to 7 until no negative entries remain in the index row.

Basis	x_1	x_2	w_1	w_2	b	check
x_2	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	3	4
w_2	1	0	$-\frac{2}{7}$	$\frac{1}{7}$	4	$\frac{34}{7}$
f	-3	0	2	0	12	11

B.O.2.

Basis	x_1	x_2	w_1	w_2	b	check
x_2	0	1	$\frac{5}{14}$	$\frac{1}{14}$	5	$\frac{45}{7}$
x_1 ~~w_2~~	1	0	$-\frac{2}{7}$	$\frac{1}{7}$	4	$\frac{34}{7}$
f	0	0	$\frac{8}{7}$	$\frac{3}{7}$	24	$\frac{179}{7}$

B.O.3.

B.O.3.

$$f + 0 \cdot x_1 + 0 \cdot x_2 + \frac{8}{7} w_1 + \frac{3}{7} w_2 = 24 \quad \Rightarrow \quad f = 24 - \frac{8}{7} w_1 - \frac{3}{7} w_2, \quad w_1, w_2 \geq 0$$

A new basic solution emerges as $x_1 = 4$, $x_2 = 5$. Since there are no negative entries in the index row, this is also the optimal solution. The optimal value for f is given in the b column, i.e. $f_{\max} = 24$.

Simplex Method (re-cap)

We have gone through the simplex in some detail by way of explanation. The solution for the problem would normally look like this:

Maximize $f = x_1 + 4x_2$ subject to the constraints

$$-x_1 + 2x_2 \le 6$$
$$5x_1 + 4x_2 \le 40$$
$$x_1, x_2 \ge 0 .$$

Entering slack variables w_1, w_2, this is written as

$$-x_1 + 2x_2 + w_1 = 6$$
$$5x_1 + 4x_2 + w_2 = 40$$
$$f - x_1 - 4x_2 = 0 .$$

Complete Simplex Table

Basis	x_1	x_2	w_1	w_2	b	check	
w_1	-1	2	1	0	6	8	
w_2	5	4	0	1	40	50	
f	-1	-4	0	0	0	-5	
w_1	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	3	4	B.O.2.
w_2	5	4	0	1	40	50	
f	-1	-4	0	0	0	-5	
x_2 ~~w_1~~	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	3	4	
w_2	7	0	-2	1	28	34	B.O.3.
f	-3	0	2	0	12	11	
x_2	$-\frac{1}{2}$	1	$\frac{1}{2}$	0	3	4	
w_2	1	0	$-\frac{2}{7}$	$\frac{1}{7}$	4	$\frac{34}{7}$	B.O.2.
f	-3	0	2	0	12	11	
x_2	0	1	$\frac{5}{14}$	$\frac{1}{14}$	5	$\frac{45}{7}$	B.O.3.
x_1 ~~w_2~~	1	0	$-\frac{2}{7}$	$\frac{1}{7}$	4	$\frac{34}{7}$	
f	0	0	$\frac{8}{7}$	$\frac{3}{7}$	24	$\frac{179}{7}$	B.O.3.

$$\therefore \quad f_{max} = 24 \quad \text{with} \quad x_1 = 4, \; x_2 = 5$$

…Simplex_Ex_1…

Container Production Optimization Problem (re-visit)

The Problem: Suppose that in producing two types of containers K and L one uses two machines M_1 and M_2. To produce a container K, one needs M_1 two minutes and M_2 four minutes. Similarly, L occupies M_1 eight minutes and M_2 four minutes. The net profit for a container K is \$29 and for L it is \$45. Determine the production plan that maximizes the net profit per hour.

Formulation: If we produce x_1 containers K and x_2 containers L per hour, the profit is $f(x_1, x_2) = 29x_1 + 45x_2$ subject to the constraints:

$$\left.\begin{aligned} 2x_1 + 8x_2 &\le 60 \\ 4x_1 + 4x_2 &\le 60 \\ x_1 \ge 0,\ x_2 &\ge 0 \end{aligned}\right\} \qquad \begin{aligned} 2x_1 + 8x_2 + w_1 &= 60 \\ 4x_1 + 4x_2 + w_2 &= 60 \\ f - 29x_1 - 45x_2 &= 0 \\ x_1, x_2, w_1, w_2 &\ge 0 \end{aligned}$$

Simplex Method…

basis	x_1	x_2	w_1	w_2	b	check
	2	8	1	0	60	71
	4	4	0	1	60	69
f	–29	–45	0	0	0	–74

$x_1, x_2, w_1, w_2 \geq 0$

basis	x_1	x_2	w_1	w_2	b	check
	0.25	1	0.125	0	7.5	8.875
	4	4	0	1	60	69
f	–29	–45	0	0	0	–74

B.O.2.

basis	x_1	x_2	w_1	w_2	b	check
	0.25	1	0.125	0	7.5	8.875
	3	0	–0.5	1	30	33.5
f	–17.75	0	5.625	0	337.5	325.375

B.O.3.

B.O.3.

Simplex Method…

$x_1, x_2, w_1, w_2 \geq 0$

basis	x_1	x_2	w_1	w_2	b	check
	0.25	1	0.125	0	7.5	8.875
	3	0	–0.5	1	30	33.5
f	–17.75	0	5.625	0	337.5	325.375

basis	x_1	x_2	w_1	w_2	b	check
	0.25	1	0.125	0	7.5	8.875
	1	0	–0.166	0.333	10	11.1666
f	–17.75	0	5.625	0	337.5	325.375

B.O.2.

basis	x_1	x_2	w_1	w_2	b	check
x_2	0	**1**	0.166	–0.083	**5**	6.083
x_1	**1**	0	–0.166	0.333	**10**	11.1666
f	0	0	2.666	5.917	515	523.582

B.O.3.

B.O.3.

…Simplex_Ex_2…

$$f + 2.666 w_1 + 5.917 w_2 = 515 \;\Rightarrow\; f = -2.666 w_1 - 5.917 w_2 + 515 \;\Rightarrow\; f_{\max} = 515$$

Simplex Method for More Variables

Many problems in real life involve more than just two variables. However, the method of computation and the key idea remain basically the same. It is an iterative process which is repeated until the index row contains no negative entry, at which point the optimal value of the objective function is attained.

We illustrate the process for solving optimization problems for more variables thru the following examples…

Example of 3 Variables

Maximize $f = 2x_1 + 6x_2 + 4x_3$ subject to the constraints

$$2x_1 + 5x_2 + 2x_3 \le 38$$
$$4x_1 + 2x_2 + 3x_3 \le 57$$
$$x_1 + 3x_2 + 5x_3 \le 57$$
$$x_1, x_2, x_3 \ge 0 .$$

Introduction of slack variables w_1, w_2, w_3 gives

$$2x_1 + 5x_2 + 2x_3 + w_1 = 38$$
$$4x_1 + 2x_2 + 3x_3 + w_2 = 57$$
$$x_1 + 3x_2 + 5x_3 + w_3 = 57$$
$$f - 2x_1 - 6x_2 - 4x_3 = 0 .$$

Basis	x_1	x_2	x_3	w_1	w_2	w_3	b	check
w_1	2	5	2	1	0	0	38	48
w_2	4	2	3	0	1	0	57	67
w_3	1	3	5	0	0	1	57	67
f	-2	-6	-4	0	0	0	0	-12
w_1	$\frac{2}{5}$	1	$\frac{2}{5}$	$\frac{1}{5}$	0	0	$\frac{38}{5}$	$\frac{48}{5}$
w_2	4	2	3	0	1	0	57	67
w_3	1	3	5	0	0	1	57	67
f	-2	-6	-4	0	0	0	0	-12
x_2 ~~w_1~~	$\frac{2}{5}$	1	$\frac{2}{5}$	$\frac{1}{5}$	0	0	$\frac{38}{5}$	$\frac{48}{5}$
w_2	$\frac{16}{5}$	0	$\frac{11}{5}$	$-\frac{2}{5}$	1	0	$\frac{209}{5}$	$\frac{239}{5}$
w_3	$-\frac{1}{5}$	0	$\frac{19}{5}$	$-\frac{3}{5}$	0	1	$\frac{171}{5}$	$\frac{191}{5}$
f	$\frac{2}{5}$	0	$-\frac{8}{5}$	$\frac{6}{5}$	0	0	$\frac{228}{5}$	$\frac{228}{5}$

B.O.2.

B.O.3.

Basis	x_1	x_2	x_3	w_1	w_2	w_3	b	check
x_2	$\frac{2}{5}$	1	$\frac{2}{5}$	$\frac{1}{5}$	0	0	$\frac{38}{5}$	$\frac{48}{5}$
w_2	$\frac{16}{5}$	0	$\frac{11}{5}$	$-\frac{2}{5}$	1	0	$\frac{209}{5}$	$\frac{239}{5}$
B.O.2. → w_3	$-\frac{1}{19}$	0	(1)	$-\frac{3}{19}$	0	$\frac{5}{19}$	9	$\frac{191}{19}$
f	$\frac{2}{5}$	0	$-\frac{8}{5}$	$\frac{6}{5}$	0	0	$\frac{228}{5}$	$\frac{228}{5}$
x_2	$\frac{8}{19}$	1	0	$\frac{5}{19}$	0	$-\frac{2}{19}$	4	$\frac{106}{19}$
w_2	$\frac{63}{19}$	0	0	$-\frac{1}{19}$	1	$-\frac{11}{19}$	22	$\frac{448}{19}$ ← **B.O.3.**
x_3 ~~w_3~~	$-\frac{1}{19}$	0	1	$-\frac{3}{19}$	0	$\frac{5}{19}$	9	$\frac{191}{19}$
f	$\frac{16}{19}$	0	0	$\frac{18}{19}$	0	$\frac{8}{19}$	60	$\frac{1172}{19}$

$$\therefore \quad f_{\max} = 60 \ \text{ with } \ x_1 = 0,\ x_2 = 4,\ x_3 = 9.$$

…Simplex_Ex_3…

$$f + \frac{16}{19}x_1 + \frac{18}{19}w_1 + \frac{8}{19}w_3 = 60 \quad \Rightarrow \quad f = 60 - \frac{16}{19}x_1 - \frac{18}{19}w_1 - \frac{8}{19}w_3, \quad x_1, w_1, w_3 \ge 0$$

Example with $\geq$ Constraint

Maximize $f = 7x_1 + 4x_2$ subject to the constraints

$$2x_1 + x_2 \leq 150$$
$$4x_1 + 3x_2 \leq 350$$
$$x_1 + x_2 \geq 80$$
$$x_1, x_2 \geq 0 .$$

$$-x_1 - x_2 \leq -80$$

Introduction of slack variables w_1, w_2, w_3 gives

$$2x_1 + x_2 + w_1 = 150$$
$$4x_1 + 3x_2 + w_2 = 350$$
$$-x_1 - x_2 + w_3 = -80$$
$$x_1, x_2, w_1, w_2, w_3 \geq 0$$

This innovative procedure was suggested by Yin Mingbao, a student taking EE2012 in Semester 1 of Academic Year 2009/2010.

Complete Simplex Table

basis	x_1	x_2	w_1	w_2	w_3	b	check
w_1	2	1	1	0	0	150	154
w_2	4	3	0	1	0	350	358
w_3	–1	–1	0	0	1	–80	–81
f	–7	–4	0	0	0	0	–11
w_1	1	0.5	0.5	0	0	75	77
w_2	4	3	0	1	0	350	358
w_3	–1	–1	0	0	1	–80	–81
f	–7	–4	0	0	0	0	–11
x_1 ~~w_1~~	1	0.5	0.5	0	0	75	77
w_2	0	1	–2	1	0	50	50
w_3	0	–0.5	0.5	0	1	–5	–4
f	0	–0.5	3.5	0	0	525	528
x_1 ~~w_1~~	1	0.5	0.5	0	0	75	77
w_2	0	1	–2	1	0	50	50
w_3	0	1	–1	0	–2	10	8
f	0	–0.5	3.5	0	0	525	528

B.O.3.

Complete Simplex Table (cont.)

basis	x_1	x_2	w_1	w_2	w_3	b	check
x_1 ~~w_1~~	1	0	1	0	1	70	73
w_2	0	0	–1	1	2	40	42
x_2 ~~w_3~~	0	1	–1	0	–2	10	8
f	0	0	3	0	–1	530	532
x_1	1	0	1	0	1	70	73
w_3 ~~w_2~~	0	0	–0.5	0.5	1	20	21
x_2	0	1	–1	0	–2	10	8
f	0	0	3	0	–1	530	532
x_1	1	0	1.5	–0.5	0	50	52
w_3	0	0	–0.5	0.5	1	20	21
x_2	0	1	–2	1	0	50	50
f	0	0	2.5	0.5	0	550	553

B.O.3.

$$\therefore \quad f_{\max} = 550 \quad \text{with} \quad x_1 = 50, \; x_2 = 50, \; w_3 = 20$$

…Simplex_Ex_4…

$$f + 2.5w_1 + 0.5w_2 = 550 \quad \Rightarrow \quad f = 550 - 2.5w_1 - 0.5w_2, \quad w_1, w_2 \ge 0$$

Another Example with ≥ Constraint

Maximize $f = 8x_1 + 4x_2$ subject to the constraints

$$2x_1 + 3x_2 \le 120$$
$$x_1 + x_2 \le 45$$
$$-3x_1 + 5x_2 \ge 25$$
$$x_1,\, x_2 \ge 0\,.$$

$$3x_1 - 5x_2 \le -25$$

Introduction of slack variables w_1, w_2, w_3 gives

$$2x_1 + 3x_2 + w_1 = 120$$
$$x_1 + x_2 + w_2 = 45$$
$$3x_1 - 5x_2 + w_3 = -25$$
$$x_1, x_2, w_1, w_2, w_3 \ge 0$$

We are to solve this problem once again using the procedure suggested by Yin Mingbao.

Simplex Table…

basis	x_1	x_2	w_1	w_2	w_3	b	check
w_1	2	3	1	0	0	120	126
w_2	1	1	0	1	0	45	48
w_3	3	–5	0	0	1	–25	–26
f	–8	–4	0	0	0	0	–12
w_1	2	3	1	0	0	120	126
w_2	1	1	0	1	0	45	48
w_3	3	–5	0	0	1	–25	–26
f	–8	–4	0	0	0	0	–12
w_1	0	1	1	–2	0	30	30
x_1 ~~w_2~~	1	1	0	1	0	45	48
w_3	0	–8	0	–3	1	–160	–170
f	0	4	0	8	0	360	372

B.O.3.

$$f + 4x_2 + 8w_2 = 360 \quad \Rightarrow \quad f = 360 - 4x_2 - 8w_2, \quad x_2, w_2 = 0 \quad \Rightarrow \quad f_{\max} = 360?$$

$$-8x_2 - 3w_2 + w_3 = -160 \quad \Rightarrow \quad w_3 = -160 < 0.$$ **This is a contradiction!**

What to do next?

The remaining control variable.

The row has problem…

basis	x_1	x_2	w_1	w_2	w_3	b	check
w_1	0	1	1	–2	0	30	30
x_1	1	1	0	1	0	45	48
w_3	0	–8	0	–3	1	–160	–170
f	0	4	0	8	0	360	372

Pivot: The intersection of the row containing w_3 and the control variable x_2

— to get rid of w_3 in the basis and obtain an explicit solution for x_2.

Simplex Table (cont.)

basis	x_1	x_2	w_1	w_2	w_3	b	check
w_1	0	1	1	−2	0	30	30
x_1	1	1	0	1	0	45	48
w_3	0	1	0	3/8	−1/8	20	21.25
f	0	4	0	8	0	360	372
w_1	0	0	1	−19/8	1/8	10	8.75
x_1	1	0	0	5/8	1/8	25	26.75
x_2	0	1	0	3/8	−1/8	20	21.25
f	0	0	0	6.5	0.5	280	287

B.O.3.

$$\therefore \quad f_{\max} = 280 \quad \text{with} \quad x_1 = 25,\ x_2 = 20,\ w_1 = 10$$

$$f + 6.5 w_2 + 0.5 w_3 = 280 \quad \Rightarrow \quad f = 280 - 6.5 w_2 - 0.5 w_3, \quad w_2, w_3 \ge 0$$

...Simplex_Ex_5...

Summary of Constrained Optimization (Simplex Method)

Given a set of linear equations and/or inequalities (constraints), the problem is to determine a solution such that a certain objective function is maximized

↓

Construct a Simplex table with slack variables

↓

Perform iteratively a series of basic operations (B.O.2 & B.O.3)

↓

The maximum value of the objective function can be determined when all its resulting coefficients are nonnegative