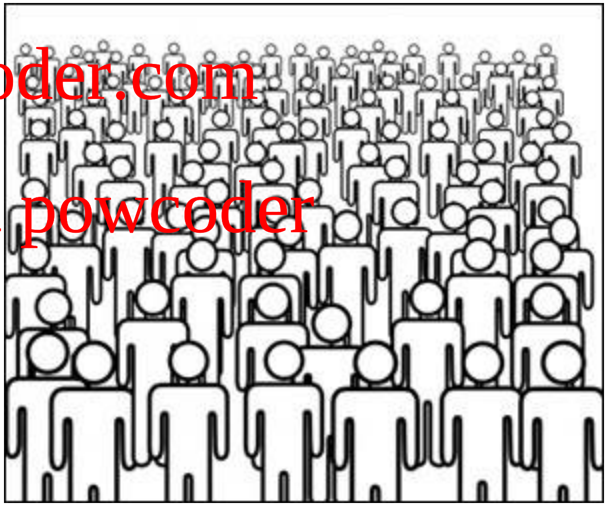
Population-based algorithms

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder



ME 564/SYS 564
Wed Oct 17, 2018
Steven Hoffenson

Goal of Week 8: To do another KKT example, and then learn some increasingly-common algorithms that use multiple points in each iteration

Recap: How to optimize

1. **Formulate** the problem

   (Weeks 1-2, 4, 9-12)

   a) Define system boundaries
   b) Develop analytical models
   c) Explore/reduce the problem space
   d) Formalize optimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}, \mathbf{p}) \\ \text{subject to} \quad & \mathbf{g}(\mathbf{x}, \mathbf{p}) \le 0 \\ & \mathbf{h}(\mathbf{x}, \mathbf{p}) = 0 \end{aligned}$$

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

TODAY

2. **Solve** the problem

   (Weeks 3, 5-8, 12)

   a) Choose the right approach/algorithm
   b) Solve (by hand, code, or software)
   c) Interpret the results
   d) Iterate if needed

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{H}(\mathbf{x}_k)]^{-1} \nabla f(\mathbf{x}_0)$$

Recap: Gradient-based methods

- **Unconstrained**
  - **FONC and SOSC** (when math is simple enough for algebra)
  - **Gradient descent** (algorithm with linear convergence)
  - **Newton method** (algorithm with hyper-linear convergence)

Assignment Project Exam Help

- **Constrained**

https://powcoder.com

  - **Reduced gradient** (with known active constraints)
  - **Generalized Reduced Gradient** (algorithm w active constr)

Add WeChat powcoder

  - **Active set strategy** (algorithm w updating set of active constr)
  - **Lagrangian** (equality or active inequality constr)
  - **KKT conditions** (with any inequality and equality constr)
  - **Quasi-Newton methods** (2^{nd}-derivative-free)
  - **SQP** (efficiently handles constraints)

Recap: Quasi-Newton and SQP

Quasi-Newton methods approximate $\mathbf{H}^{-1}$ to simplify math

1. Begin with $\mathbf{x}_0$ and some assumed $\mathbf{H}_0^{-1}$.
2. For iteration k, set $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{H}_k^{-1} \nabla f(\mathbf{x}_k)$.
3. Compute an update matrix $\hat{\mathbf{H}}_k^{-1}$ as some function of $[\nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)]$, $[\mathbf{x}_{k+1} - \mathbf{x}_k]$, and $\mathbf{H}_k^{-1}$.
4. Update inverse Hessian approximation: $\mathbf{H}_{k+1}^{-1} = \mathbf{H}_k^{-1} + \hat{\mathbf{H}}_k^{-1}$.

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Sequential Quadratic Programming (SQP) algorithm solves a subproblem for the step size and direction $\mathbf{s}_k$, then moves as $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{s}_k$

$$\begin{aligned}
&\underset{\mathbf{s}_k}{\text{minimize}} && q(\mathbf{s}_k) = f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)^T \mathbf{s}_k + \frac{1}{2} \mathbf{s}_k^T \nabla_{xx}^2 \mathscr{L}(\mathbf{x}_k, \lambda_k, \mu_k) \mathbf{s}_k \\
&\text{where} && \mathscr{L}(\mathbf{x}_k, \lambda_k, \mu_k) = f(\mathbf{x}_k) - \lambda^T \mathbf{h}(\mathbf{x}_k) - \mu^T \mathbf{g}(\mathbf{x}_k) \\
&\text{subject to} && \mathbf{g}(\mathbf{x}_k) + \nabla \mathbf{g}(\mathbf{x}_k)^T \mathbf{s}_k \le \mathbf{0} \\
& && \mathbf{h}(\mathbf{x}_k) + \nabla \mathbf{h}(\mathbf{x}_k)^T \mathbf{s}_k = \mathbf{0}
\end{aligned}$$

Example: KKT conditions (5.14)

$$\min \ f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + 40x_1 + 20x_2 - 3000$$

s.t. $g_1(\mathbf{x}) = x_1 - 50 \geq 0$

$g_2(\mathbf{x}) = x_1 + x_2 - 100 \geq 0$

$g_3(\mathbf{x}) = x_1 + x_2 + x_3 - 150 \geq 0$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Recall the KKT conditions:

1. $\boldsymbol{\nabla}_{\mathbf{x}}\mathcal{L} = \boldsymbol{\nabla}\mathrm{f}(\mathbf{x}^*) + \boldsymbol{\lambda}^T\boldsymbol{\nabla}\mathbf{h}(\mathbf{x}^*) + \boldsymbol{\mu}^T\boldsymbol{\nabla}\mathbf{g}(\mathbf{x}^*) = \mathbf{0}^T$

2. $\mathbf{h}(\mathbf{x}^*) = \mathbf{0}, \mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$

3. $\boldsymbol{\lambda} \neq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0}$

4. $\boldsymbol{\mu}^T\mathbf{g}(\mathbf{x}^*) = \mathbf{0}^T$

Example: KKT conditions (5.14)

$$\begin{aligned} \min\ & f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + 40x_1 + 20x_2 - 3000 \\ \text{s.t.}\ & g_1(\mathbf{x}) = x_1 - 50 \geq 0 \\ & g_2(\mathbf{x}) = x_1 + x_2 - 100 \geq 0 \\ & g_3(\mathbf{x}) = x_1 + x_2 + x_3 - 150 \geq 0 \end{aligned}$$

The constraints all need to be multiplied by -1 to become negative null!

Assignment Project Exam Help

$$\nabla_{\mathbf{x}}\mathcal{L} = \nabla f(\mathbf{x}^*) + \boldsymbol{\lambda}^{\mathrm{T}}\nabla \mathbf{h}(\mathbf{x}^*) + \boldsymbol{\mu}^{\mathrm{T}}\nabla \mathbf{g}(\mathbf{x}^*) = \mathbf{0}^{\mathrm{T}}$$

https://powcoder.com

Add WeChat powcoder

$$\nabla f(\mathbf{x}^*) = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \partial f / \partial x_3 \end{bmatrix} \qquad \nabla \mathbf{g}(\mathbf{x}^*) = \begin{bmatrix} \partial g_1 / \partial x_1 & \partial g_1 / \partial x_2 & \partial g_1 / \partial x_3 \\ \partial g_2 / \partial x_1 & \partial g_2 / \partial x_2 & \partial g_2 / \partial x_3 \\ \partial g_3 / \partial x_1 & \partial g_3 / \partial x_2 & \partial g_3 / \partial x_3 \end{bmatrix}$$

$$\nabla_{\mathbf{x}}\mathcal{L} = \begin{bmatrix} 2x_1 + 40 \\ 2x_2 + 20 \\ 2x_3 \end{bmatrix} + \mathbf{0}^{\mathrm{T}} + \begin{bmatrix} \mu_1 & \mu_2 & \mu_3 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ -1 & -1 & 0 \\ -1 & -1 & -1 \end{bmatrix} = \mathbf{0}^{\mathrm{T}}$$

Example: KKT conditions (5.14)

$$\nabla_{\mathbf{x}}\mathcal{L} = \begin{bmatrix} 2x_1 + 40 \\ 2x_2 + 20 \\ 2x_3 \end{bmatrix} + \mathbf{0}^{\mathbf{T}} + \begin{bmatrix} \mu_1 & \mu_2 & \mu_3 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ -1 & -1 & 0 \\ -1 & -1 & -1 \end{bmatrix} = \mathbf{0}^{\mathbf{T}}$$

Assignment Project Exam Help

$$2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0$$
$$2x_2 + 20 - \mu_2 - \mu_3 = 0$$
$$2x_3 - \mu_3 = 0$$

https://powcoder.com

3 equations,
6 unknowns

Add WeChat powcoder

Recall the KKT conditions:

1. $\nabla_{\mathbf{x}}\mathcal{L} = \mathbf{0}^{\mathbf{T}}$
2. $\mathbf{h}(\mathbf{x}^*) = \mathbf{0}$, $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$
3. $\boldsymbol{\lambda} \neq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$
4. $\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x}^*) = \mathbf{0}^{\mathbf{T}}$

$$\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x}^*) = \mathbf{0}^{\mathbf{T}}$$

$$\begin{bmatrix} \mu_1 & \mu_2 & \mu_3 \end{bmatrix} \begin{bmatrix} -x_1 + 50 \\ -x_1 - x_2 + 100 \\ -x_1 - x_2 - x_3 + 150 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

3 more equations!

Example: KKT conditions (5.14)

$$2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0$$
$$2x_2 + 20 - \mu_2 - \mu_3 = 0$$
$$2x_3 - \mu_3 = 0$$
$$\mu_1(-x_1 + 50) = 0$$
$$\mu_2(-x_1 - x_2 + 100) = 0$$
$$\mu_3(-x_1 - x_2 - x_3 + 150) = 0$$

6 equations,
6 unknowns

Assignment Project Exam Help

https://powcoder.com

How do we solve?

Add WeChat powcoder

From the lower 3 equations, there are multiple possibilities in each: Either $\mu_i = 0$, or the expression in parentheses $= 0$.

So, we will have to examine 8 different scenarios for this system of equations, and check each solution against the remaining KKT conditions: $\cancel{\mathbf{h}(\mathbf{x}^*) = \mathbf{0}}$, $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, $\cancel{\boldsymbol{\lambda} \neq \mathbf{0}}$, $\boldsymbol{\mu} \geq \mathbf{0}$.

We don't need these, since there are no equality constraints in this problem

Example: KKT conditions (5.14)

Scenario 1: $\mu_1 = 0, \mu_2 = 0, \mu_3 = 0$

$$
\boxed{
\begin{aligned}
&2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\
&2x_2 + 20 - \mu_2 - \mu_3 = 0 \\
&2x_3 - \mu_3 = 0 \\
&\mu_1(-x_1 + 50) = 0 \\
&\mu_2(-x_1 - x_2 + 100) = 0 \\
&\mu_3(-x_1 - x_2 - x_3 + 150) = 0
\end{aligned}
}
$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

$$
\begin{aligned}
&2x_1 + 40 = 0 \\
&2x_2 + 20 = 0 \\
&2x_3 = 0 \\
&0 = 0 \\
&0 = 0 \\
&0 = 0
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
&x_1 = -20 \\
&x_2 = -10 \\
&x_3 = 0 \\
&\mu_1 = 0 \\
&\mu_2 = 0 \\
&\mu_3 = 0
\end{aligned}
\quad \Rightarrow
$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0}$

$$g_1 = 20 + 50 \nleq 0$$

Since g_1 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 2: $\mu_1 \neq 0, \mu_2 = 0, \mu_3 = 0$

$$
\boxed{\begin{aligned}
&2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0\\
&2x_2 + 20 - \mu_2 - \mu_3 = 0\\
&2x_3 - \mu_3 = 0\\
&\mu_1(-x_1 + 50) = 0\\
&\mu_2(-x_1 - x_2 + 100) = 0\\
&\mu_3(-x_1 - x_2 - x_3 + 150) = 0
\end{aligned}}
$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

$$
\begin{aligned}
&2x_1 + 40 - \mu_1 = 0\\
&2x_2 + 20 = 0\\
&2x_3 = 0\\
&-x_1 + 50 = 0\\
&0 = 0\\
&0 = 0
\end{aligned}
$$

$$
\begin{aligned}
&x_1 = 50\\
&x_2 = -10\\
&x_3 = 0\\
&\mu_1 = 140\\
&\mu_2 = 0\\
&\mu_3 = 0
\end{aligned}
$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$

$$g_2 = -50 + 10 + 100 \nleq 0$$

Since g_2 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 3: $\mu_1 = 0, \mu_2 \neq 0, \mu_3 = 0$

$$
\boxed{\begin{aligned}
&2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\
&2x_2 + 20 - \mu_2 - \mu_3 = 0 \\
&2x_3 - \mu_3 = 0 \\
&\mu_1(-x_1 + 50) = 0 \\
&\mu_2(-x_1 - x_2 + 100) = 0 \\
&\mu_3(-x_1 - x_2 - x_3 + 150) = 0
\end{aligned}}
$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

$$
\begin{aligned}
&2x_1 + 40 - \mu_2 = 0 \\
&2x_2 + 20 - \mu_2 = 0 \\
&2x_3 = 0 \\
&0 = 0 \\
&-x_1 - x_2 + 100 = 0 \\
&0 = 0
\end{aligned}
$$

$$
\begin{aligned}
&x_1 = 45 \\
&x_2 = 55 \\
&x_3 = 0 \\
&\mu_1 = 0 \\
&\mu_2 = 130 \\
&\mu_3 = 0
\end{aligned}
$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0}$

$$g_1 = -45 + 50 \nleq 0$$

Since g_1 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 4: $\mu_1 = 0, \mu_2 = 0, \mu_3 \neq 0$

$$
\begin{aligned}
&2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\
&2x_2 + 20 - \mu_2 - \mu_3 = 0 \\
&2x_3 - \mu_3 = 0 \\
&\mu_1(-x_1 + 50) = 0 \\
&\mu_2(-x_1 - x_2 + 100) = 0 \\
&\mu_3(-x_1 - x_2 - x_3 + 150) = 0
\end{aligned}
$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

$$
\begin{aligned}
&2x_1 + 40 - \mu_3 = 0 \\
&2x_2 + 20 - \mu_3 = 0 \\
&2x_3 - \mu_3 = 0 \\
&0 = 0 \\
&0 = 0 \\
&-x_1 - x_2 - x_3 + 150 = 0
\end{aligned}
$$

$$
\begin{aligned}
&x_1 = 40 \\
&x_2 = 50 \\
&x_3 = 60 \\
&\mu_1 = 0 \\
&\mu_2 = 0 \\
&\mu_3 = 120
\end{aligned}
$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$

$$g_1 = -40 + 50 \nleq 0$$

Since g_1 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 5: $\mu_1 \neq 0, \mu_2 \neq 0, \mu_3 = 0$

$$\boxed{\begin{aligned} &2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\ &2x_2 + 20 - \mu_2 - \mu_3 = 0 \\ &2x_3 - \mu_3 = 0 \\ &\mu_1(-x_1 + 50) = 0 \\ &\mu_2(-x_1 - x_2 + 100) = 0 \\ &\mu_3(-x_1 - x_2 - x_3 + 150) = 0 \end{aligned}}$$

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

$$\begin{aligned} &2x_1 + 40 - \mu_1 - \mu_2 = 0 \\ &2x_2 + 20 - \mu_2 = 0 \\ &2x_3 = 0 \\ &-x_1 + 50 = 0 \\ &-x_1 - x_2 + 100 = 0 \\ &0 = 0 \end{aligned}$$

$$\begin{aligned} x_1 &= 50 \\ x_2 &= 50 \\ x_3 &= 0 \\ \mu_1 &= 20 \\ \mu_2 &= 120 \\ \mu_3 &= 0 \end{aligned}$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$

$g_3 = -50 - 50 - 0 + 150 \nleq 0$

Since g_3 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 6: $\mu_1 \neq 0, \mu_2 = 0, \mu_3 \neq 0$

$$
\boxed{\begin{aligned}
&2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0\\
&2x_2 + 20 - \mu_2 - \mu_3 = 0\\
&2x_3 - \mu_3 = 0\\
&\mu_1(-x_1 + 50) = 0\\
&\mu_2(-x_1 - x_2 + 100) = 0\\
&\mu_3(-x_1 - x_2 - x_3 + 150) = 0
\end{aligned}}
$$

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

$$
\begin{aligned}
&2x_1 + 40 - \mu_1 - \mu_3 = 0\\
&2x_2 + 20 - \mu_3 = 0\\
&2x_3 - \mu_3 = 0\\
&-x_1 + 50 = 0\\
&0 = 0\\
&-x_1 - x_2 - x_3 + 150 = 0
\end{aligned}
$$

$$
\begin{aligned}
x_1 &= 50\\
x_2 &= 45\\
x_3 &= 55\\
\mu_1 &= 30\\
\mu_2 &= 0\\
\mu_3 &= 110
\end{aligned}
$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$

$$g_2 = -50 - 45 + 100 \nleq 0$$

Since g_2 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 7: $\mu_1 = 0, \mu_2 \neq 0, \mu_3 \neq 0$

$$\boxed{\begin{aligned} &2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\ &2x_2 + 20 - \mu_2 - \mu_3 = 0 \\ &2x_3 - \mu_3 = 0 \\ &\mu_1(-x_1 + 50) = 0 \\ &\mu_2(-x_1 - x_2 + 100) = 0 \\ &\mu_3(-x_1 - x_2 - x_3 + 150) = 0 \end{aligned}}$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

$$\begin{aligned} &2x_1 + 40 - \mu_2 - \mu_3 = 0 \\ &2x_2 + 20 - \mu_2 - \mu_3 = 0 \\ &2x_3 - \mu_3 = 0 \\ &0 = 0 \\ &-x_1 - x_2 + 100 = 0 \\ &-x_1 - x_2 - x_3 + 150 = 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_1 &= 45 \\ x_2 &= 55 \\ x_3 &= 50 \\ \mu_1 &= 0 \\ \mu_2 &= 30 \\ \mu_3 &= 100 \end{aligned}$$

$\Rightarrow$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$

$$g_1 = -45 + 50 \nleq 0$$

Since g_1 is violated, this is not a KKT point.

Example: KKT conditions (5.14)

Scenario 8: $\mu_1 \neq 0, \mu_2 \neq 0, \mu_3 \neq 0$

$$\boxed{\begin{aligned} &2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\ &2x_2 + 20 - \mu_2 - \mu_3 = 0 \\ &2x_3 - \mu_3 = 0 \\ &\mu_1(-x_1 + 50) = 0 \\ &\mu_2(-x_1 - x_2 + 100) = 0 \\ &\mu_3(-x_1 - x_2 - x_3 + 150) = 0 \end{aligned}}$$

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

$$\begin{aligned} &2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 = 0 \\ &2x_2 + 20 - \mu_2 - \mu_3 = 0 \\ &2x_3 - \mu_3 = 0 \\ &-x_1 + 50 = 0 \\ &-x_1 - x_2 + 100 = 0 \\ &-x_1 - x_2 - x_3 + 150 = 0 \end{aligned} \Rightarrow \begin{aligned} x_1 &= 50 \\ x_2 &= 50 \\ x_3 &= 50 \\ \mu_1 &= 20 \\ \mu_2 &= 20 \\ \mu_3 &= 100 \end{aligned}$$

Check: $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0}$

All of these conditions hold, so $(50, 50, 50)^T$ is a KKT point!

Example: KKT conditions (5.14)

$$
\begin{aligned}
\min \;\; & f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + 40x_1 + 20x_2 - 3000 \\
\text{s.t.} \;\; & g_1(\mathbf{x}) = x_1 - 50 \geq 0 \\
& g_2(\mathbf{x}) = x_1 + x_2 - 100 \geq 0 \\
& g_3(\mathbf{x}) = x_1 + x_2 + x_3 - 150 \geq 0
\end{aligned}
$$

Assignment Project Exam Help

Across the 8 scenarios, the only KKT point is $\mathbf{x}^* = (50, 50, 50)^T$, $f^* = 7500$

https://powcoder.com

Now we can test the SOSC:

Add WeChat powcoder

$$
\nabla_{\mathbf{x}}\mathcal{L} = \begin{bmatrix} 2x_1 + 40 - \mu_1 - \mu_2 - \mu_3 \\ 2x_2 + 20 - \mu_2 - \mu_3 \\ 2x_3 - \mu_3 \end{bmatrix}
$$

This is the **analytical** approach to solving constrained problems. We also have **algorithmic** approaches like SQP

$$
\mathcal{L}_{xx} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}
$$

← This is pos. def. everywhere!

Therefore, $\mathbf{x}^*$ is a **global minimizer**.

Summary: Gradient-based algorithms

- Gradient descent
- Newton method
- Generalized Reduced Gradient (GRG)
- Active set strategy
- Quasi-Newton methods
- Sequential Quadratic Programming (SQP)

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

$$\nabla f$$

Gradient-free approaches

- Approximation models
- Pattern search (e.g., Hooke-Jeeves, Nelder-Meade)
- Space-filling search (e.g., DIRECT)
- Random search (e.g., Simulated Annealing)
- Linear Programming (e.g., Simplex)
- Genetic/evolutionary algorithms
- Particle swarm
- Ant colony

We covered some of these in Week 3

These are population-based algorithms

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Approximation models

Use gradient-based methods even though your functions are not differentiable

1. Approximate derivatives with finite differences

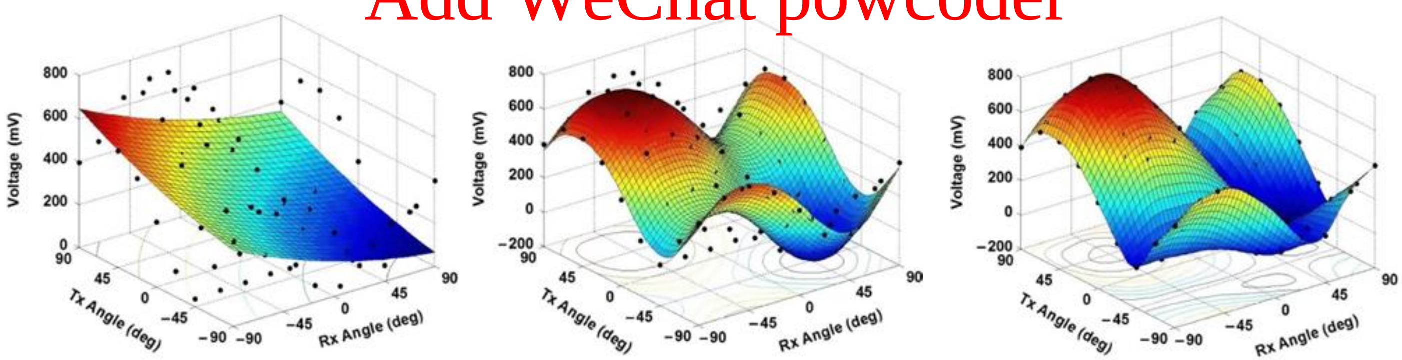
Assignment Project Exam Help

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \text{ for some small } h$$

https://powcoder.com

2. Metamodels: Sample points & fit a function

Add WeChat powcoder



Figures from:

Zamani, A., Mirabadi, A., and Schmid, F. (2012), "Applying metamodeling techniques in the design and optimization of train wheel detector", *Sensor Review*, 32 (4), 327 – 336.

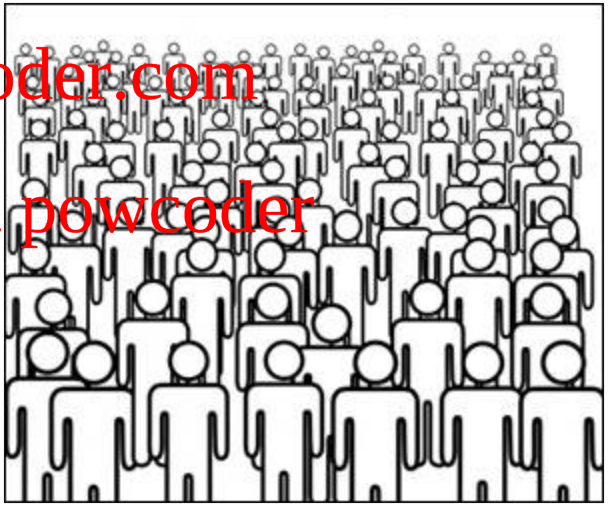
Population-based algorithms

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

ME 564/SYS 564
Wed Oct 17, 2018
Steven Hoffenson

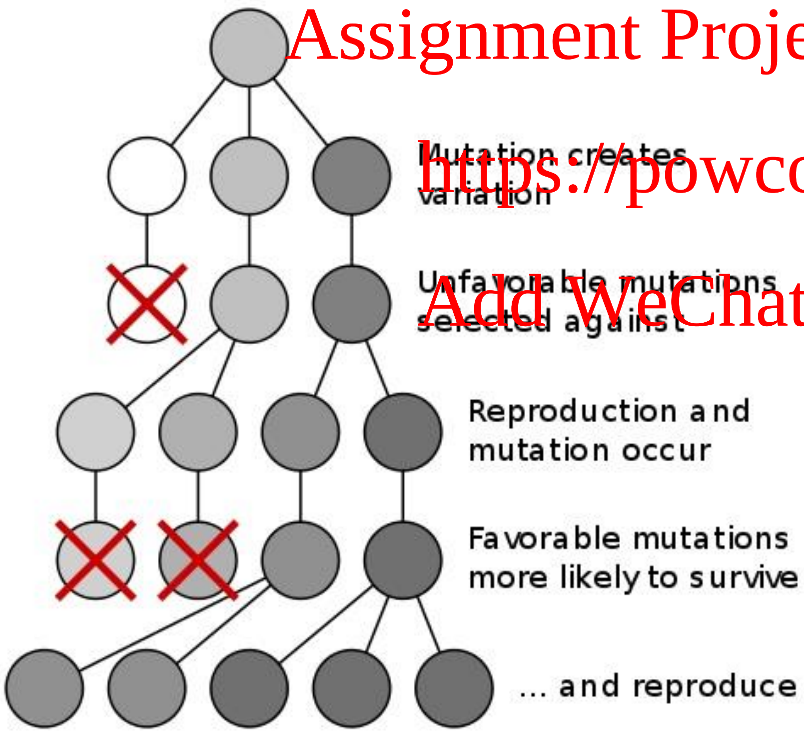


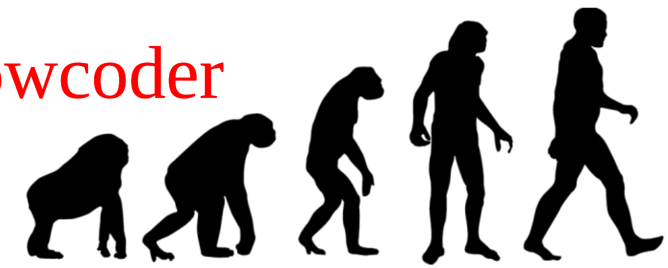
Goal of Week 8: To learn some increasingly-common algorithms that use multiple points in each iteration, and practice using them

Genetic algorithms – overview

- A.k.a., evolutionary algorithms
- Mimic gene selection and "survival of the fittest"

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder





Genetic algorithms – steps

1. Start with a random population (set) of inputs
2. Select the best among population to be “parents”
3. Use parents to spawn a new generation through crossovers and mutations
4. Repeat 2-3 with more “generations” until satisfied

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

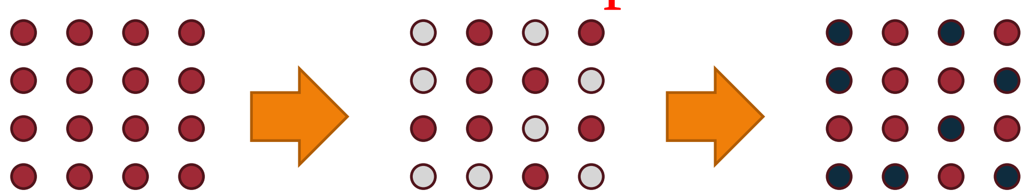


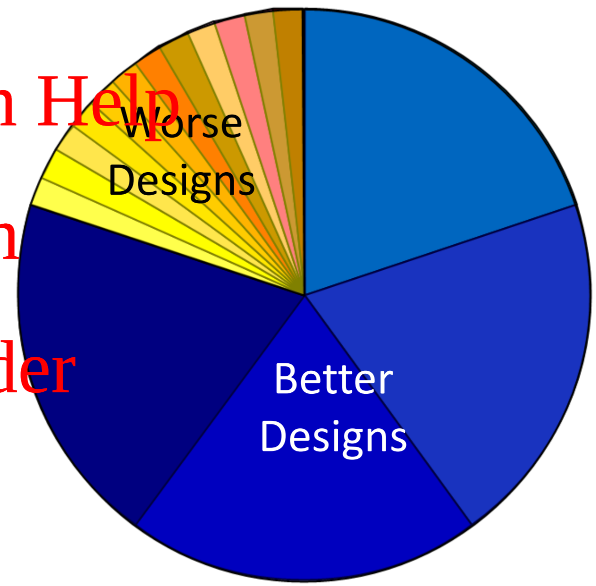
figure from John Whitefoot, 2010 lecture

Genetic algorithms – parent selection

There are different ways to select parents from a population:

- **Elitism:** pick only the absolute best points
- **Roulette wheel:** better designs have a better chance of being chosen (pictured)
- **Tournament:** segment the population randomly, and choose the best in each segment

Different algorithms use different strategies.



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Genetic algorithms – spawning

Spawning new generations may be done using a combination of:

1. **Survivors:** the best simply join the new generation
2. **Crossovers:** combinations of traits from 2 parents create a child
3. **Mutations:** traits from a parent randomly change

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

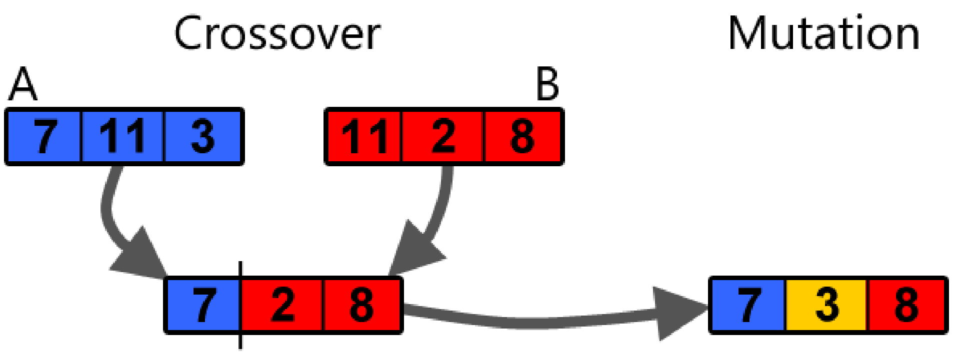
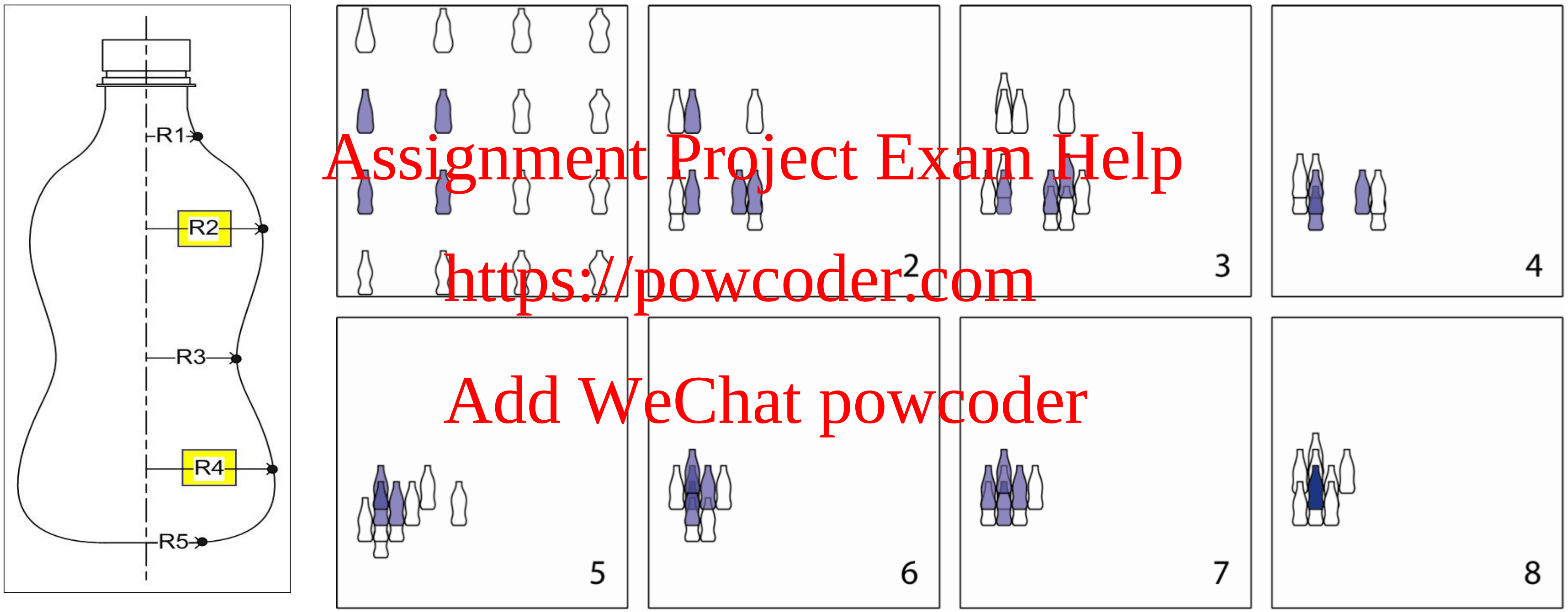


Image credit: http://www.klopfenstein.net/lorenz.aspx/genetic-algorithms (accessed 2013)

Example: *Interactive* Genetic Algorithm

Which bottle shape do you prefer?



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

The interactive part means that humans do the "parent selection" portion in each iteration of the GA. In this case, it usually converged to a Coca-Cola bottle shape.

Dr. Jarod Kelly, University of Michigan

MATLAB – ga

There is a genetic algorithm function ‘ga’ that can handle objectives and constraints in a similar way to fmincon

Assignment Project Exam Help

```
[xopt,fopt]=ga(OBJFUN,NVARS,A,b,Aeq,beq,lb,ub,NONLCON)
```

https://powcoder.com

Note the difference:

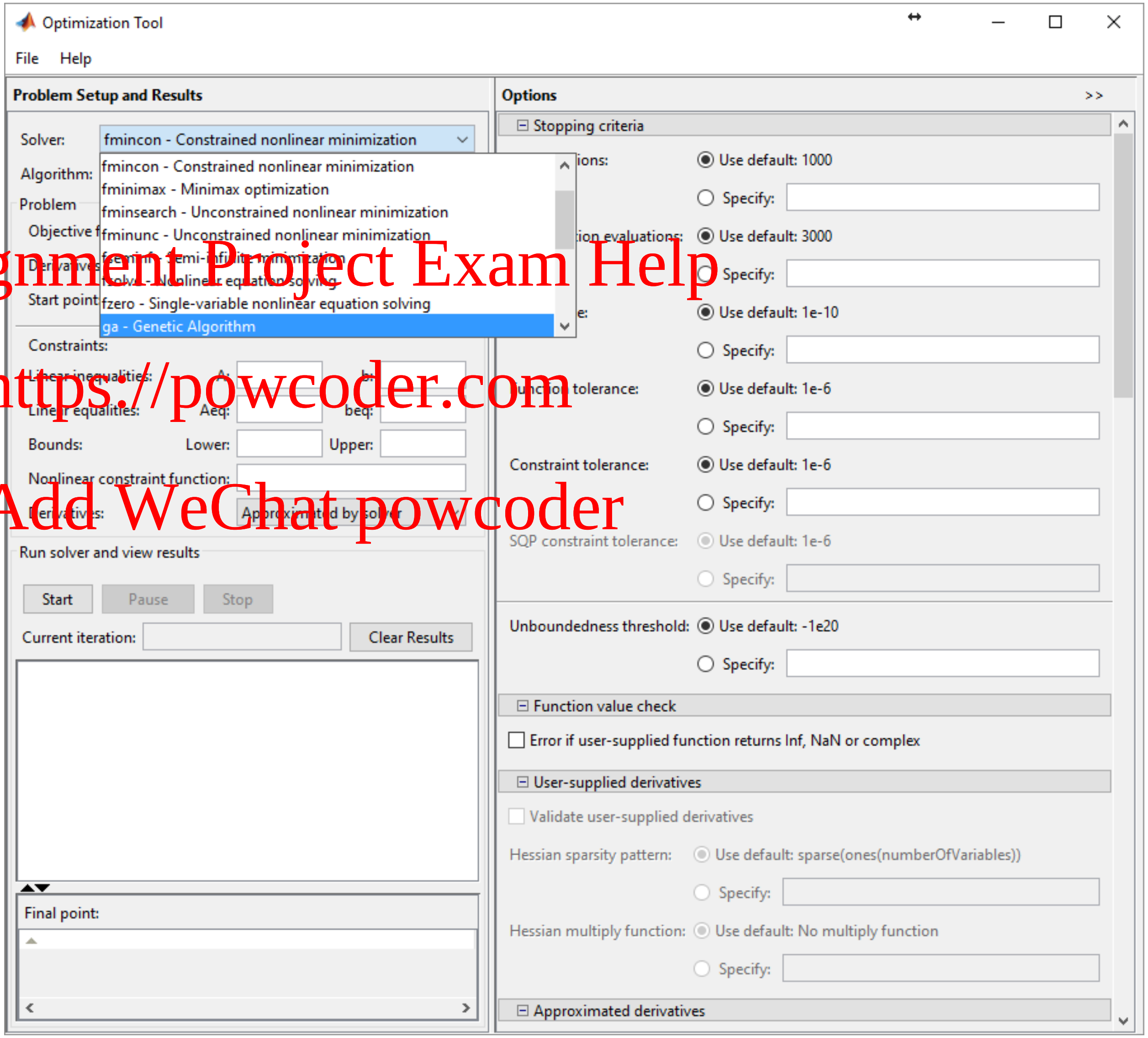
Add WeChat powcoder

- fmincon needs a start point
- ga just needs the # of variables

Note: This function requires the “Global optimization toolbox” to be installed.

MATLAB – optimtool

MATLAB currently has an interactive tool ‘optimtool’ that can guide you through optimizing with different options



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Note: This may disappear in future MATLAB releases ☹

Particle swarm optimization

1. Start with some random set of points (particles) with directions/velocities within the input space
2. At each iteration, update each particle’s position and velocity based on the particle’s best previous position and the best positions of its near neighbors

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

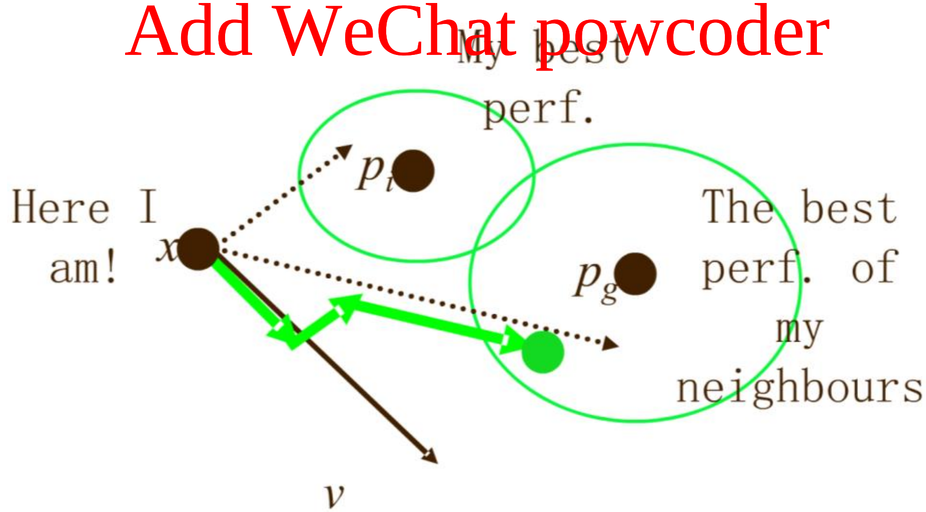
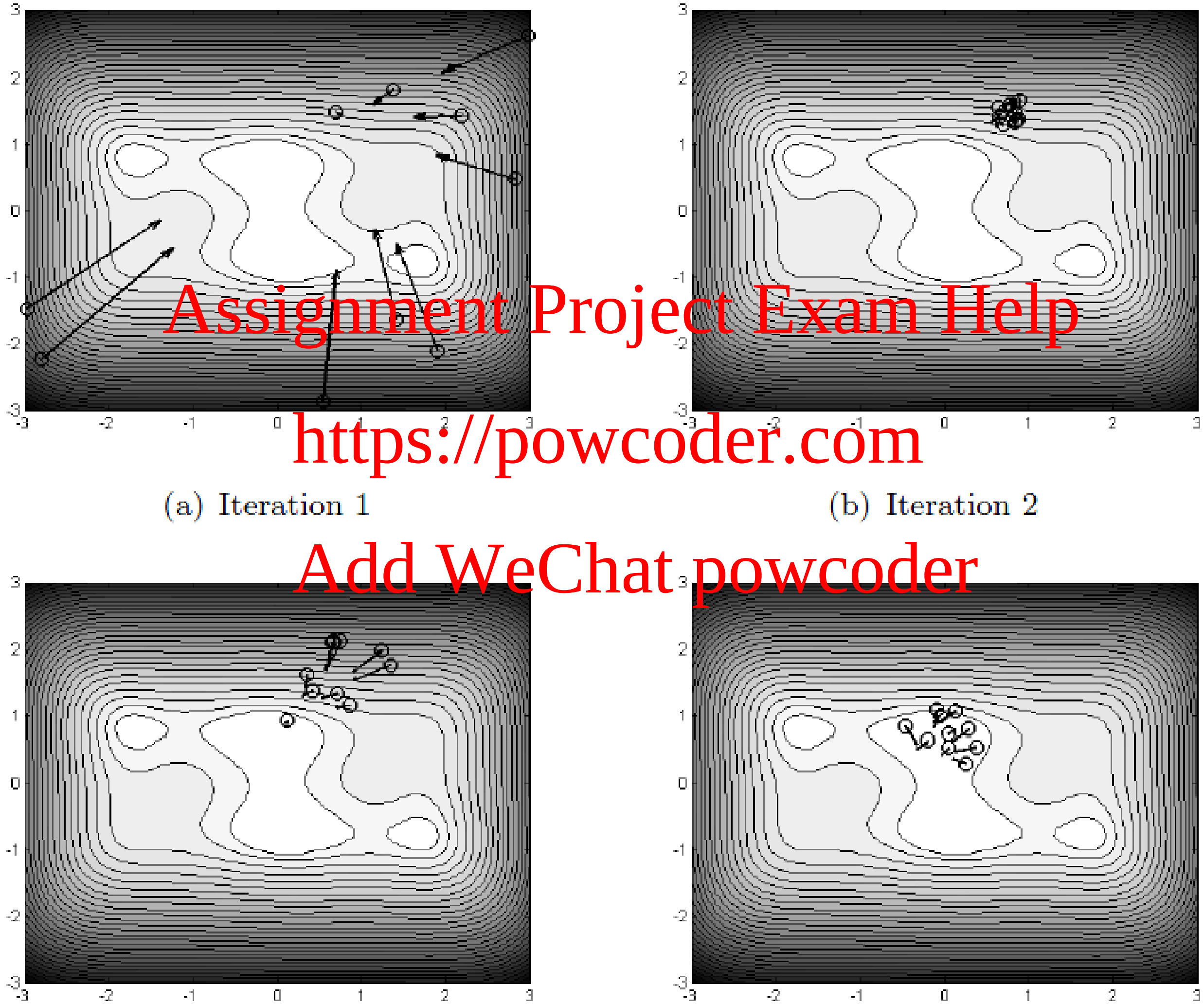


Figure from a presentation by Maurice Clerc at (accessed 2013):
http://www.docstoc.com/docs/76778278/Particle-Swarm-Optimization-mini-tutorial-(PowerPoint)

Particle swarm optimization



(a) Iteration 1 (b) Iteration 2

(c) Iteration 3 (d) Iteration 4

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

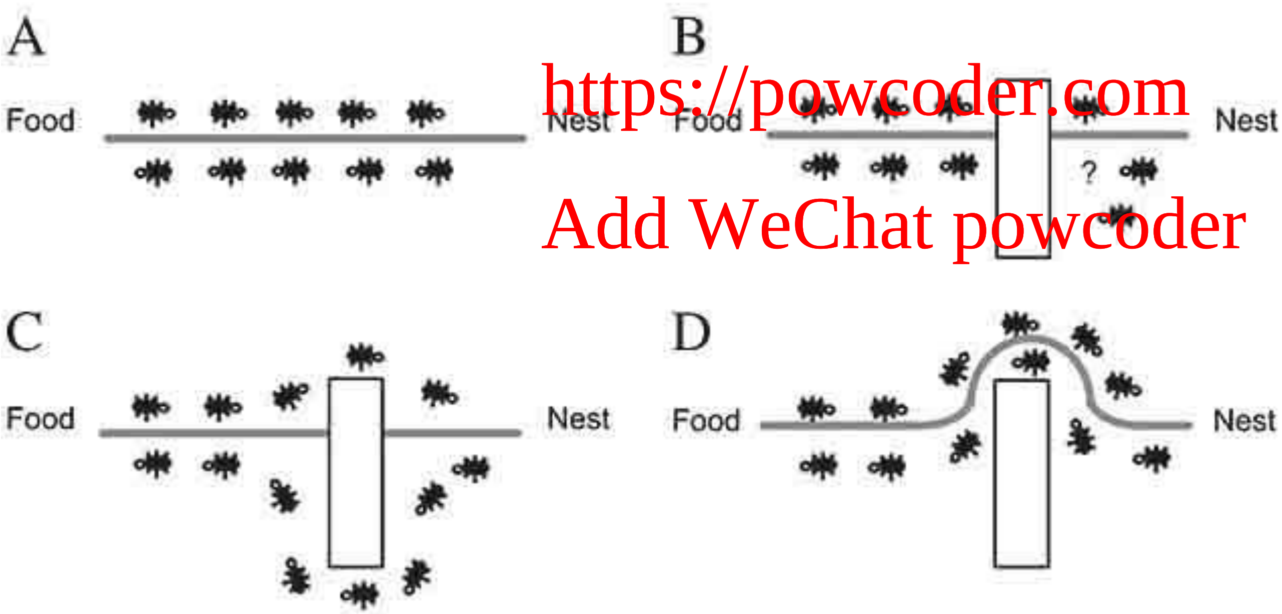
Ant colony optimization

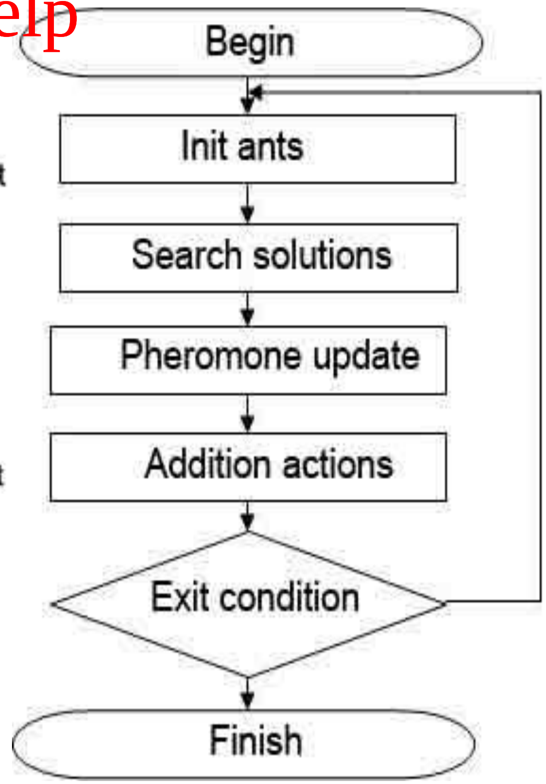
Adjustments depend on a combination of randomness and pheromones (which tell what has been tried and true), which evaporate over time

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder





Figures from http://masters.donntu.edu.ua/2011/fknt/el_khatib/diss/indexe.htm (accessed 2013)

Non-continuous/discrete variables

What to do when we have **non-continuous** variables?

Examples:

- Number of rotors on a drone (must be integer)
- Number of cylinders in an engine (must be integer)
- Material choice (there are a few options with different specific properties)

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Strategies:

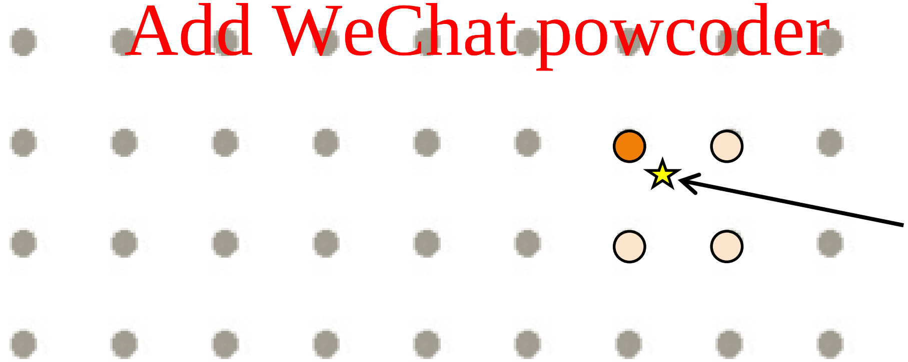
1. Treat them as continuous, and then pick nearby point
2. Parametric optimization
3. Integer programming

Treat discrete variables as continuous

The **easiest** way to handle non-continuous variables (material properties, countable things) is to treat them as continuous:

1. Optimize pretending that they are continuous
2. Choose the discrete value closest to the optimizer (or evaluate all surrounding points and pick the best)

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder



Optimizer with continuous assumption

This often does not work well with non-convex functions!

Parametric optimization

When the number of discrete choices is small, you can do ***parametric optimization***:

1. List out all discrete variable combinations
Assignment Project Exam Help
2. Optimize to find the best solution under each of the scenarios in (1)
https://powcoder.com
3. Choose the best from all the solutions in (2)
Add WeChat powcoder

This works well when there is a small number of discrete choices and the functions are quick to evaluate

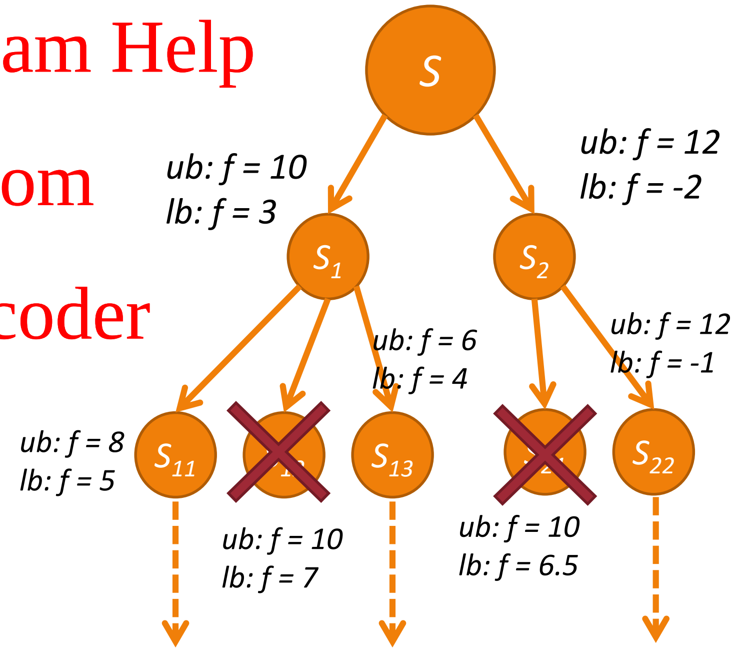
Discrete/integer programming

When you have a large number of non-continuous variables…

Branch-and-bound algorithm

1. Branch: Split the space of candidate solutions
2. Bound: Compute upper and lower bounds on each subspace (i.e., figure out the highest and lowest possible objective function values for each group)
3. If the lower bound for one group is higher than the upper bound for another, eliminate it
4. Repeat 1-3 until a single solution

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

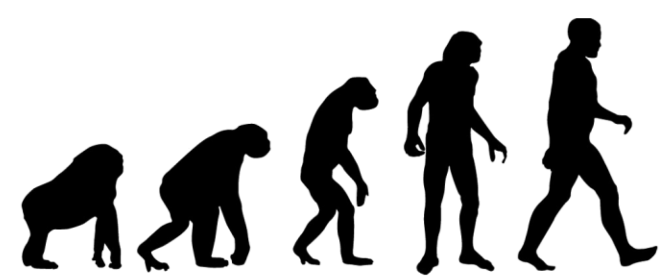


Recap: Gradient-free algorithms

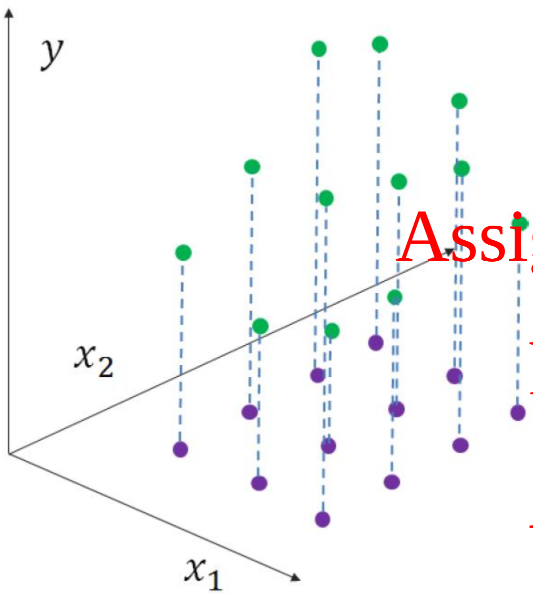
- Approximation models
- Pattern search (e.g., Hooke-Jeeves)
- Space-filling search (e.g., DIRECT)
- Random search (e.g., Simulated Annealing)
- Linear Programming (e.g., Simplex)
- Genetic/evolutionary algorithms
- Particle swarm
- Ant colony

Week 3

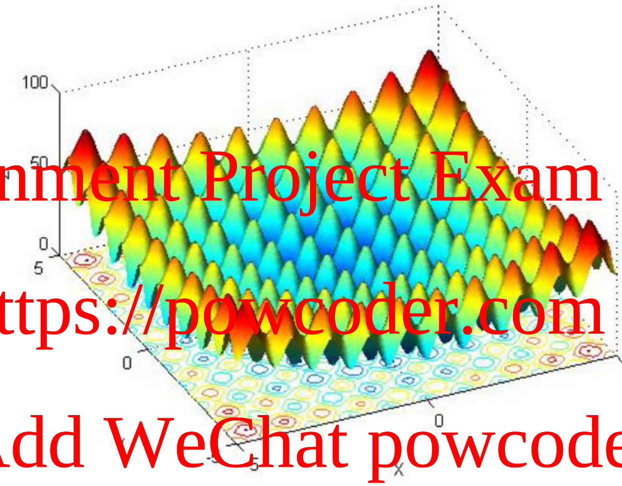
Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder



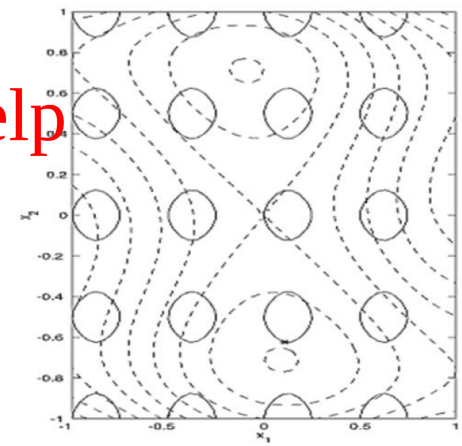
When to use gradient-free methods



No Gradient



Too Many Local Minima



Complicated Design Space

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

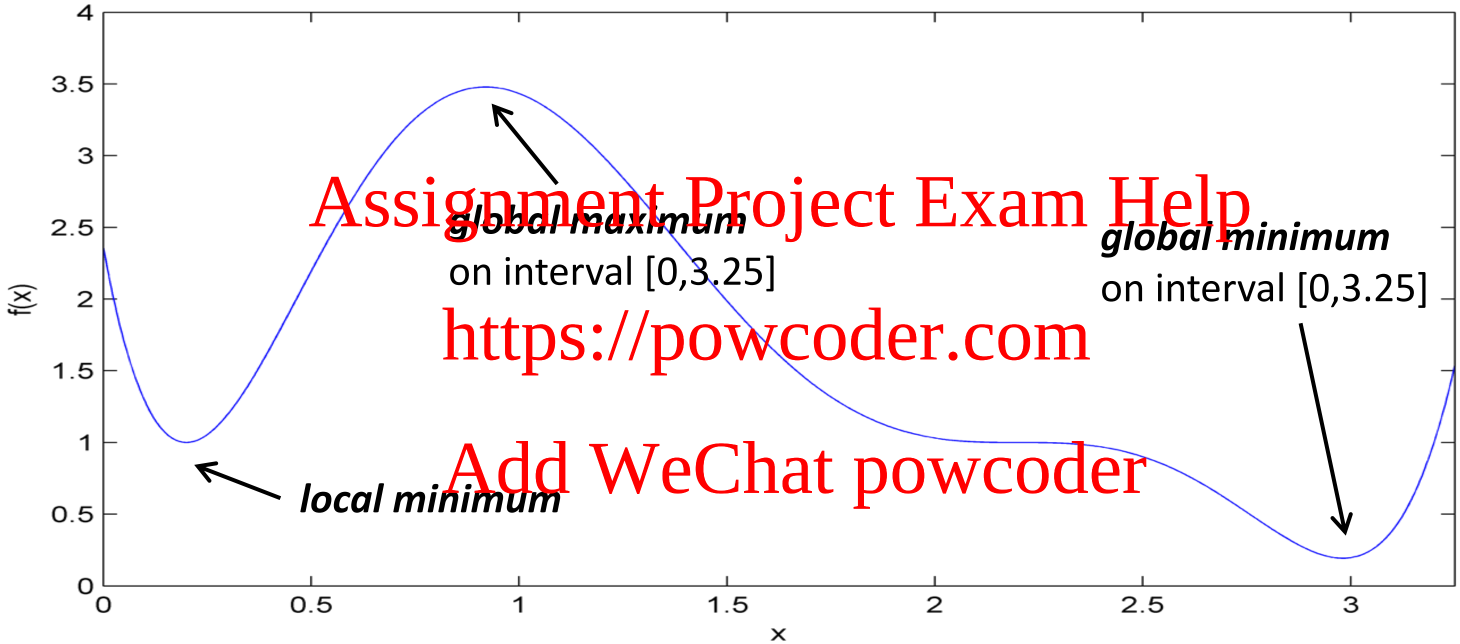
Cons of gradient-free methods

- Usually slower to converge
  - These require many more function evaluations
  - The search directions are not always efficient
- No optimality guarantee
  - There is no analogy to FONC (grad = 0) and SOSC (Hessian is positive definite)
  - Convergence must be measured by changes to f and x
- Many parameters to tune
- Constraint handling is imperfect
  - Cannot use Lagrangian
  - Must use penalty or barrier
- Stochastic methods are not repeatable

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

A note on global optimization

Most algorithms seek local optima



Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

To find global solutions, try:

1. Performing optimization with multiple start points
2. Using global algorithms (e.g., genetic algorithms & particle swarm)

Recap: How to optimize

1. **Formulate** the problem
   a) Define system boundaries
   b) Develop analytical models
   c) Explore/reduce the problem space
   d) Formalize optimization problem

(Weeks 1-2, 4, 9-12)

$$\underset{\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x}, \mathbf{p})$$
$$\text{subject to} \quad \mathbf{g}(\mathbf{x}, \mathbf{p}) \leq 0$$
$$\mathbf{h}(\mathbf{x}, \mathbf{p}) = 0$$

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

TODAY

2. **Solve** the problem
   a) Choose the right approach/algorithm
   b) Solve (by hand, code, or software)
   c) Interpret the results
   d) Iterate if needed

(Weeks 3, 5-8, 12)

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{H}(\mathbf{x}_k)]^{-1} \nabla f(\mathbf{x}_0)$$

Summary of optimization approaches

- Mathematical/analytical
  - Elimination of constraints using monotonicity analysis
  - Finding stationary points and their nature with optimality conditions
- Linear programming: Simplex
- Nonlinear gradient-based methods
  - Gradient descent
  - Newton method
  - Generalized Reduced Gradient (GRG)
  - Active set strategies
  - Quasi-Newton methods
  - Sequential quadratic programming (SQP)
- Nonlinear gradient-free methods
  - Approximation
  - Pattern search
  - Space-filling search
  - Random search
  - Genetic/evolutionary algorithms (GAs/EAs)
  - Particle swarm & Ant colony
- Integer programming: Branch-and-bound

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

When to use what?

- If the **math is simple** enough, try solving by hand (monotonicity analysis, optimality conditions: FONC/SOSC or KKT)
- If you have a **linear** problem, Simplex LP is efficient
- If the **functions are way too slow** to evaluate, create a meta-model
- If you have a **convex and differentiable** problem, you can't beat the efficiency and accuracy of a gradient-based algorithm (e.g., SQP, GRG)
- If you have a **convex, non-differentiable** problem, pattern-search or random search algorithms are usually effective
- If you have **non-continuous variables**, you should either:
  - Solve parametrically (i.e., solve separately for each discrete value)
  - Use branch-and-bound techniques
- With tricky problems (**non-linear, non-convex**) with fast functions, try:
  - Convex search methods with multiple start points
  - Gradient-free algorithms like GA's and other bio-inspired searches

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder

Acknowledgements

- Some of this material came from Chapter 7 of the textbook, *Principles of Optimal Design*
- Some of these slides and examples came from Dr. John Whitefoot, Dr. Alex Burnap, Dr. Yi Ren, and Dr. Michael Kokkolaras at the University of Michigan

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Announcements

- HW4 is posted, due in 13 days, Tuesday at noon
- Project progress reports are due a week from Sunday
- Please do the mid-semester survey! It’s totally anonymous and will help me improve the course

  https://goo.gl/forms/hR7fN7luoFUtJOKU2

Assignment Project Exam Help
https://powcoder.com
Add WeChat powcoder