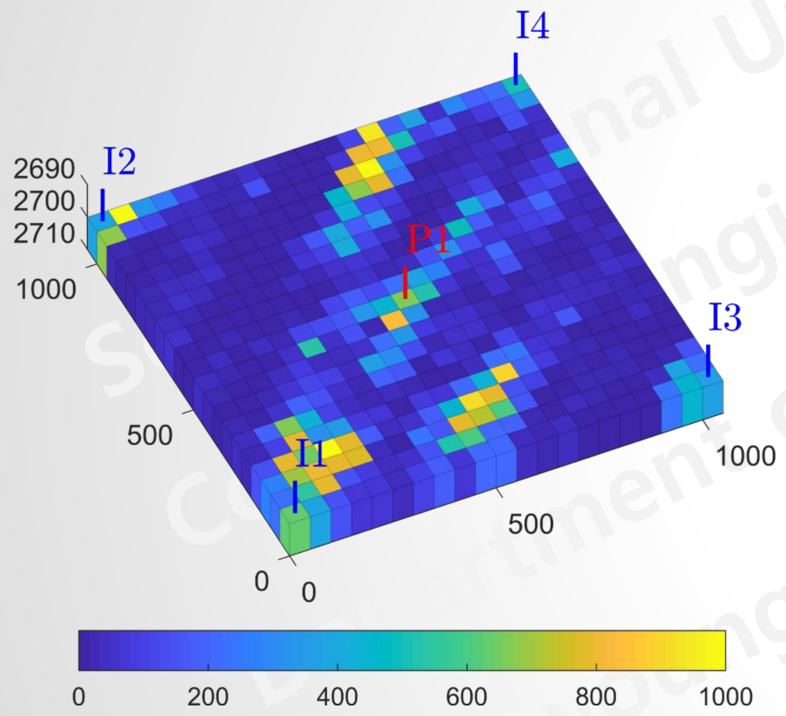
Robust Optimization and Stochastic Gradient

Hoonyoung Jeong
Department of Energy Resources Engineering
Seoul National University

ENGINEERING
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Introduction



Horizontal permeability (md)
21 by 21 by 1
Oil and water

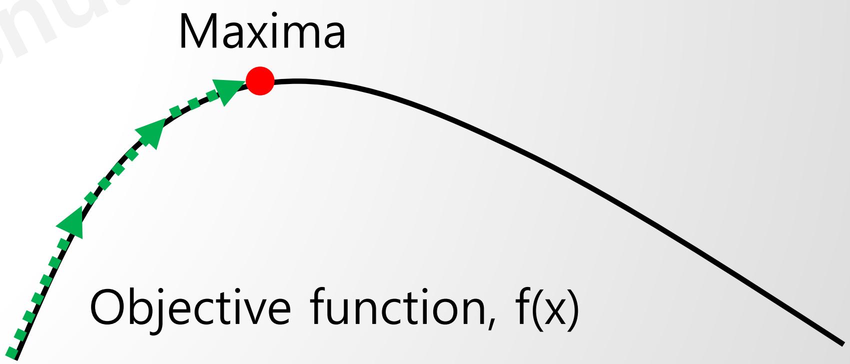
- Optimizing q_w at I1 – I4
- Optimal q_w at I1 – I4?
  - ✓Maximizing NPV

$$NPV = discount(Oil\ prod \cdot Oil\ price$$
$$- Water\ prod \cdot Water\ treatment\ unit\ cost$$
$$- Water\ inj \cdot Water\ injection\ unit\ cost)$$

  - ✓Gradient-based optimization

Gradient-based Optimization

- Gradient
  - ✓First order derivatives
  - ✓$x = [q_1 \quad q_2 \quad q_3 \quad q_4]^T, NPV = f(x)$
  - ✓$\nabla f(x) = \left[\frac{\partial f}{\partial q_1} \quad \frac{\partial f}{\partial q_2} \quad \frac{\partial f}{\partial q_3} \quad \frac{\partial f}{\partial q_4}\right]^T$
  - ✓$\nabla f(x)$: direction at x maximizing f
- How to find optimal x using a gradient
  - ✓Change x along $\nabla f(x)$
  - ✓Nonlinear
    - ❑Move a little along $\nabla f(x)$, and repeat many times
    - ❑$x_{k+1} = x_k + \lambda_k \nabla f(x_k)$

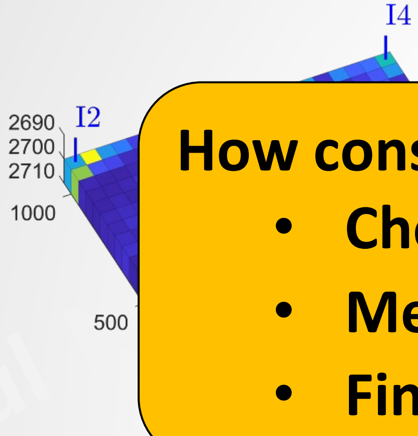


What if Reservoir Model Parameters are Uncertain?

k and ϕ are known at only I1-I4 and P1

How consider reservoir uncertainty in optimization?

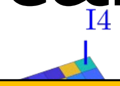
- **Choose x_{opt} with maximum NPV?**
- **Mean of x_{opt} for 5 models?**
- **Find x maximizing the mean of NPV of 5 models!**



En1

$x_{opt} = [47 \quad 45 \quad 45 \quad 33]^T \; bbl/day$

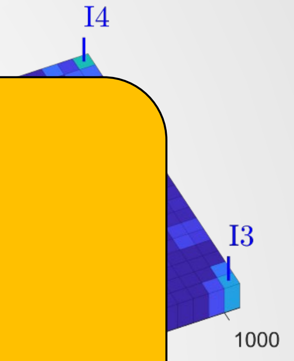
NPV=9.44 MM$



En2

$x_{opt} = [43 \quad 53 \quad 53 \quad 22]^T \; bbl/day$

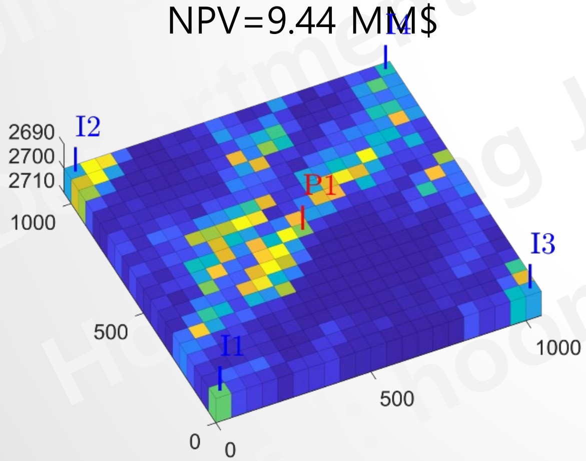
NPV=9.48 MM$



En3

$x_{opt} = [46 \quad 44 \quad 42 \quad 37]^T \; bbl/day$

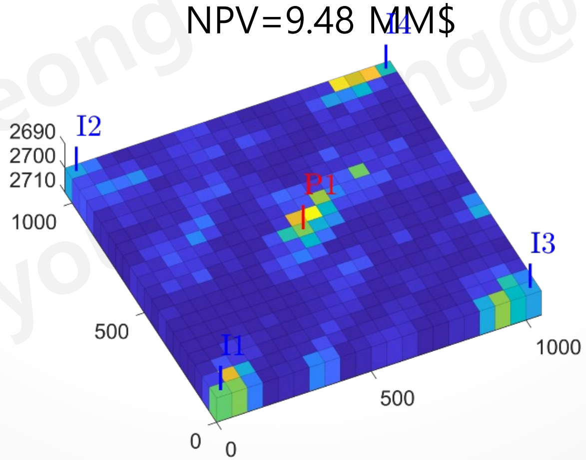
NPV=9.66 MM$



En4

$x_{opt} = [24 \quad 39 \quad 92 \quad 31]^T \; bbl/day$

NPV=9.57 MM$



En5

$x_{opt} = [39 \quad 46 \quad 45 \quad 30]^T \; bbl/day$

NPV=8.81 MM$

- Key factors for optimization
  - ✓ Pressure maintenance at P1
  - ✓ Water breakthrough time
- Well connected INJ and PRO: smaller q
- Less connected INJ and PRO: larger q
- I1 and I4: overall smaller because of the correlation direction

Robust Optimization

- Suggested to consider reservoir uncertainty on 2009
- Find an optimal solution
  - ✓Maximizing
    - ❑Objective function of a single model (X)
    - ❑Mean of objective function values of possible models (O)
    - ❑$f = \frac{1}{N_e}\sum_{m=1}^{N_e} NPV(m_i, x)$
- Too expensive
  - ✓Ex)
    - ❑100 simulation runs for a single model
    - ❑10,000 simulation runs for 100 models

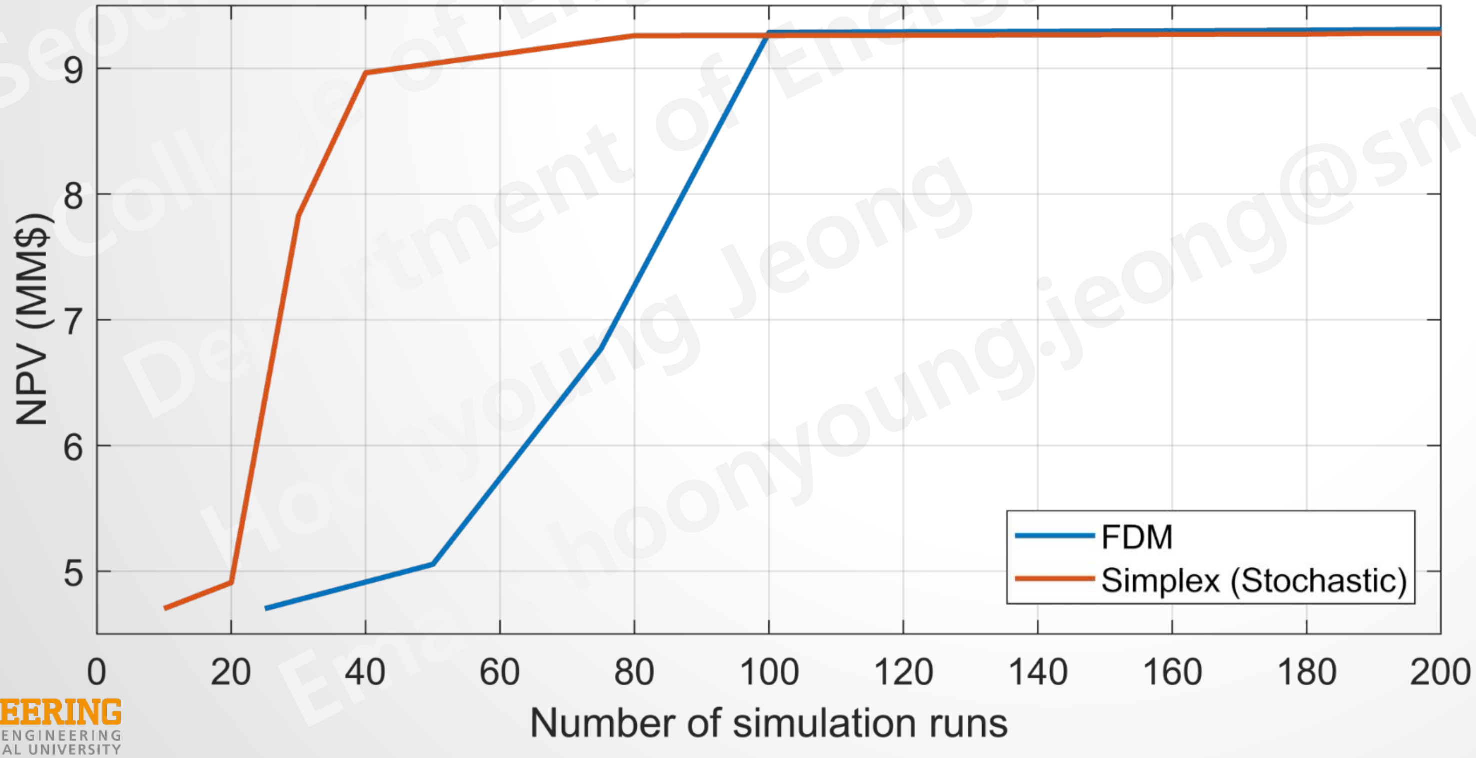
Robust Optimization using Stochastic Gradient

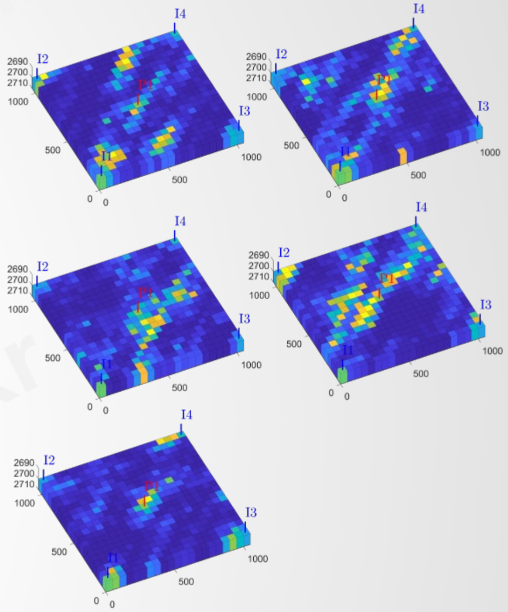
$$f = \frac{1}{N_e}\sum_{m=1}^{N_e} NPV(m_i, x)\ \ where\ N_e = 5$$

	Numerical calculation of a gradient	
	Common (FDM)	Stochastic (StoSAG)
How to approximate	Give 1% perturbations, then compute sensitivities	Give random perturbations simultaneously, then compute average sensitivities
Formula	$\nabla f(x) = \begin{bmatrix}\frac{\partial f}{\partial q_1} & \frac{\partial f}{\partial q_2} & \frac{\partial f}{\partial q_3} & \frac{\partial f}{\partial q_4}\end{bmatrix}^T \approx \begin{bmatrix}\frac{f(q_1+0.01q_1,q_2,q_3,q_4)-f(q_1,q_2,q_3,q_4)}{0.01q_1} \\ \frac{f(q_1,q_2+0.01q_2,q_3,q_4)-f(q_1,q_2,q_3,q_4)}{0.01q_2} \\ \frac{f(q_1,q_2,q_3+0.01q_3,q_4)-f(q_1,q_2,q_3,q_4)}{0.01q_3} \\ \frac{f(q_1,q_2,q_3,q_4+0.01q_4)-f(q_1,q_2,q_3,q_4)}{0.01q_4}\end{bmatrix}$	$\nabla f(x) = \frac{1}{N_e}\sum_{i=1}^{N_e}\frac{1}{N_p}\sum_{j=1}^{N_p}(\hat{\boldsymbol{x}}_{k,i,j} - \boldsymbol{x}_k)\left(NPV(\mathbf{m}_i, \hat{\boldsymbol{x}}_{k,i,j}) - NPV(\mathbf{m}_i, \boldsymbol{x}_k)\right)$
Cost	# of sim. runs = (4+1) * 5 = 25	# of sim. runs = (1+ # of perturbations) * # of models = 10
Accuracy & Cost	Acurate, but Expensive	Less accurate, but cheaper

Comparison of Computational Cost

- $f = \frac{1}{N_e}\sum_{m=1}^{N_e} NPV(m_i, x)\ \ where\ N_e = 5$
- The simplex gradient method might converge to the maxima faster than FDM, but not all the time





Why Stochastic Gradient?

- Many initial solutions should be attempted to find a global optima because gradient-based optimization methods are likely to seek local optima
- Too high computation cost is consumed to compute a gradient using FDM
- It is computationally demanding to attempt many initial solutions if FDM is used to compute a gradient
- Rather than accurate, but costly FDM gradient, what about less accurate, but cheap stochastic gradient? Attempting many initial solutions is less costly

Stochastic Gradient is Always Better than FDM?

- No, stochastic gradients is more advantageous than FDM
  - ✓When # of ensemble models is large
  - ✓When the size of x is large
  - ✓When many initial solutions should be attempted