Dual Approach for Solving The Global Minimum of The Double Well Potential Problem

Gang-Xuan Lin

Department of Mathematics, National Cheng Kung University, Tainan, Taiwan

24 May 2012

Outline

1. Introduction

Outline

Outline

Outline

Outline

Introduction

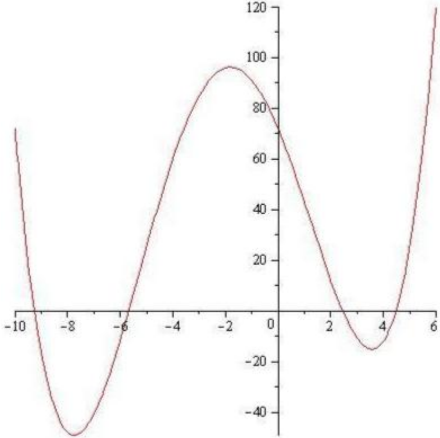
In this talk, we propose a new model which minimizes a special type of multi-variate polynomial of degree 4 as follows:

$$\min_{x\in\mathbb{R}^n} \quad \frac{1}{2}\left(\frac{1}{2}\|Bx-c\|^2-d\right)^2+\frac{1}{2}x^TAx-f^Tx. \tag{1}$$

$A \in S_{n\times n}$, $0 \neq B \in M_{m\times n}$, $c \in R^m$, $f \in R^n$, $d \in R$, whereas x is a $n \times 1$ state variable.

Introduction

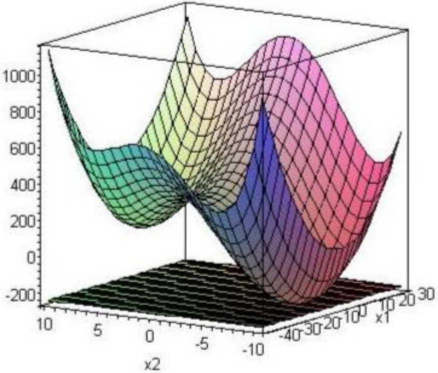
- By selecting the problem data appropriately, the (DWP) problem consists of some typical examples.
- The simplest example for $n = 1$ is:



- There are two local energy wells separated by a barrier.

Introduction

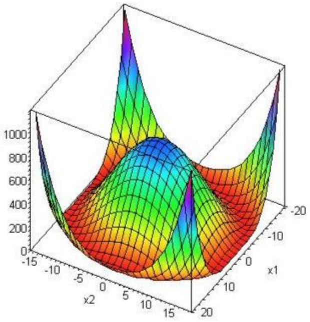
- A higher dimensional analogy is:



- The barrier is not a local maximum but a saddle point.

Introduction

- Another example for $n = 2$ is:



Mexican hat

- This interesting example is called the Mexican hat potential.
- It forms a ring-shaped of infinitely many global minima with a unique local maximum in the center.

Introduction

Due to the common feature in these illustrative examples, the (DWP) model is referred to as the *double well potential problem*.

Introduction

- Our motivation to investigate the (DWP) problem came from a numerical approximation to the generalized Ginzburg-Landau functionals.
- The functionals often describe the total energy of a ferroelectric system such as the ion-molecule reactions. [1]
- Some applications of the Ginzburg-Landau functionals can be found in solid mechanics and quantum mechanics. [2] [3]

---

[1] J. I. Brauman, Some historical background on the double-well potential model. *Journal of Mass Spectrometry, 30, 1649–1651, 1995.*

[2] D.Y. Gao, H. Yu, Multi-scale modelling and canonical dual finite element method in phase transitions of solids, International Journal of Solids and Structures, 45, 3660-3673, 2008.

[3] A. Heuer and U. Haeberlen, The dynamics of hydrogens in double well potentials: The transition of the jump rate from the low temperature quantum-mechanical to the high temperature activated regime. *J. Chem. Phys., 15: (6), 4201-4214, 1991.*

Introduction

- The mathematical formula of the generalized Ginzburg-Landau functionals takes the following general form [4] [5]

$$I^{\alpha}(\mu) = \int_{\Omega} [\frac{1}{n} \|\nabla \mu(x)\|^{n} + \frac{\alpha}{2}(\frac{1}{2}\|\mu(x)\|^{2} - \beta)^{2}] dx$$

- $\Omega \subset R^n$, α, β are positive material constants, $\mu : \Omega \longrightarrow R^q$ is a smooth vector-valued (field) function describing the phase (order) of the system.

[4] D.Y. Gao, H. Yu, Multi-scale modelling and canonical dual finite element method in phase transitions of solids, International Journal of Solids and Structures, 45, 3660-3673, 2008.

[5] R. L. Jerrard, Lower bounds for generalized Ginzburg-Landau functionals. *SIAM J. Math. Anal., 30: (4), 721-746, 1999.*

Introduction

$$I^{\alpha}(\mu) = \int_{\Omega} [\frac{1}{n} \|\nabla \mu(x)\|^{n} + \frac{\alpha}{2}(\frac{1}{2}\|\mu(x)\|^{2} - \beta)^{2}] dx$$

- The second term of the functional is indeed the double-well potential in the integral form.
- When α is sufficiently large (so that the second term dominates) and if the trace of μ on $\partial\Omega$ is a function of non-zero Brouwer degree, $I^{\alpha}(\mu)$ is bounded from below by $\ln \alpha$ [6].
- Directly minimizing $I^{\alpha}(\mu)$ over any reasonable functional space is, in general, very difficult.
- So that only the lower bound is estimated.

[6] R. L. Jerrard, Lower bounds for generalized Ginzburg-Landau functionals. *SIAM J. Math. Anal., 30: (4), 721-746, 1999.*

Introduction

$$I^{\alpha}(\mu) = \int_{\Omega} [\frac{1}{n}\|\nabla\mu(x)\|^{n} + \frac{\alpha}{2}(\frac{1}{2}\|\mu(x)\|^{2} - \beta)^{2}]dx$$

- The second term of the functional is indeed the double-well potential in the integral form.
- When α is sufficiently large (so that the second term dominates) and if the trace of μ on $\partial\Omega$ is a function of non-zero Brouwer degree, $I^{\alpha}(\mu)$ is bounded from below by $\ln\alpha$ [6].
- Directly minimizing $I^{\alpha}(\mu)$ over any reasonable functional space is, in general, very difficult.
- So that only the lower bound is estimated.

[6] R. L. Jerrard, Lower bounds for generalized Ginzburg-Landau functionals. *SIAM J. Math. Anal., 30: (4), 721-746, 1999.*

Introduction

- We therefore look into the discrete version of $I^{\alpha}(\mu)$ and it naturally leads to a special case of (DWP).
- To illustrate how $I^{\alpha}(\mu)$ can be discretized into (DWP), we work out an example with $n = 2$, $q = 1$.
- $\Omega = \Omega_x \times \Omega_y = [0, 1] \times [0, 1]$.
- $\{0 = x_1, x_2, \ldots, x_{n+1} = 1\}$ be the uniform grid of Ω_x.
- $\{0 = y_1, y_2, \ldots, y_{m+1} = 1\}$ be the uniform grid of Ω_y.
- Define a $(n+1) \times (m+1)$ vector e by

$$\begin{aligned} e = \ & [e_{1,1}, e_{2,1}, \cdots, e_{n+1,1}, e_{1,2}, e_{2,2}, \cdots, e_{n+1,2}, \\ & \cdots, e_{1,m+1}, e_{2,m+1}, \cdots, e_{n+1,m+1}]^T, \end{aligned}$$

  where $e_{i,j} = \mu(x_i, y_j)$.

Introduction

- We can approximate $\nabla\mu$ by the first order difference.
- Approximate

$$I^{\alpha}(\mu)=\int_{\Omega}[\frac{1}{2}\|\nabla\mu(x)\|^2+\frac{\alpha}{2}(\frac{1}{2}\|\mu(x)\|^2-\beta)^2]dx$$

by the Riemann sum so that a discrete version of $I^{\alpha}(\mu)$ becomes

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{2}|(\frac{e_{i+1,j}-e_{i,j}}{\frac{1}{n}})^2+(\frac{e_{i,j+1}-e_{i,j}}{\frac{1}{m}})^2|\cdot\frac{1}{n}\frac{1}{m}+\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\alpha}{2}(\frac{1}{2}e_{i,j}^2-\beta)^2\cdot\frac{1}{m}\frac{1}{n}.$$

Introduction

- The generalized Ginzburg-Landau functional $I^{\alpha}(u)$ has an estimated upper bound by

$$\frac{\alpha}{8mn}(\|e\|^2)^2 + \frac{1}{2}e^T(\frac{n}{m}\sum_{i\in T}\mathcal{B}_i + \frac{m}{n}\sum_{i\in T}\mathcal{C}_i - \frac{\alpha\beta}{mn}I)e + \frac{\alpha\beta^2}{2}. \qquad (2)$$

- $\mathcal{B}_i = diag(0_{i-1}, E, 0_{(n+1)(m+1)-i-1})$, $E = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$.
- $\mathcal{C}_i$ is a $(n+1)(m+1) \times (n+1)(m+1)$ matrix with the (i,i) and $(i+(n+1), i+(n+1))$ components being 1; $(i, i+(n+1))$ and $(i+(n+1), i)$ components being -1; and 0 elsewhere.
- $T = \{1,2,3,\cdots,(n+1)m\}\backslash\{(n+1), 2(n+1), 3(n+1), \cdots, m(n+1)\}$.
- (2) is of the form (DWP) with $x = e; B = (\frac{\alpha}{mn})^{\frac{1}{4}}I$; $A = (\frac{n}{m}\sum_{i\in T}\mathcal{B}_i + \frac{m}{n}\sum_{i\in T}\mathcal{C}_i - \frac{\alpha\beta}{mn}I)$; $c = 0$; $d = 0$; and $f = 0$.

Introduction

- The proposed (DWP) problem is thus well formulated.
- In this talk, we shall focus on solving the global minimum of (DWP).

Space Reduction and Format Setting

- The (DWP) problem

$$\min \quad \frac{1}{2}\left(\frac{1}{2}\|Bx-c\|^2-d\right)^2+\frac{1}{2}x^TAx-f^Tx.$$

- We introduce a new variable t and reduce (DWP) into the following quadratic programming with a single quadratic equality constraint, called (QP1QC):

$$\begin{array}{lll} \min & P(t,x)=\frac{1}{2}t^2+\frac{1}{2}x^TAx-f^Tx & \\ s.t. & t=\frac{1}{2}\|Bx-c\|^2-d & \\ & \quad =\frac{1}{2}x^TB^TBx-c^TBx+\frac{1}{2}c^Tc-d. & \end{array} \tag{3}$$

Space Reduction and Format Setting

- We know how to solve (QP1QC) based on the results in [7] [8]
- (QP1QC) is posed with a quadratic equality constraint whereas the results derived in Xing et al., and Feng et al. were meant for the inequality constraint.

[7] W. Xing, S.C. Fang, D.Y. Gao, R.L. Sheu and L. Zhang, Canonical Dual Solutions to the Quadratic Programming over a Quadratic Constraint, submitted.

[8] J.M. Feng, G.X. Lin, R.L. Sheu and Y. Xia, Duality and solutions for quadratic programming over single non-homogeneous quadratic constraint, to appear in *Journal of Global Optimization.*

Space Reduction and Format Setting

- $B^TB \succeq 0$.
- If B^TB singular, let U be a matrix form by basis of $N(B)$, we extend U to a nonsingular $[U, V]$.
- For every x, $x = Uy + Vz$, $y \in R^r, z \in R^{n-r}$.
- (QP1QC) can be formulated as

$$\begin{cases} \min\limits_{t,y,z} & \frac{1}{2}t^2 + \frac{1}{2}y^TA_{uu}y + \frac{1}{2}z^TA_{vv}z + y^TA_{uv}z - f^TUy - f^TVz \\ \text{s.t.} & \frac{1}{2}z^TB_{vv}z - t - (B^Tc)^TVz - (d - \frac{1}{2}c^Tc) = 0, \end{cases}$$

$A_{uu} = U^TAU, A_{vv} = V^TAV, A_{uv} = U^TAV = A_{vu}^T$ and $B_{uu} = U^TB^TBU$, $B_{vv} = V^TB^TBV, B_{uv} = U^TB^TBV = B_{vu}^T$.

- $B_{vv} \succ 0$.

Space Reduction and Format Setting

$$\begin{cases} \min\limits_{t,y,z} & \frac{1}{2}t^2 + \frac{1}{2}y^T A_{uu} y + \frac{1}{2} z^T A_{vv} z + y^T A_{uv} z - f^T U y - f^T V z \\ \text{s.t.} & \frac{1}{2} z^T B_{vv} z - t - (B^T c)^T V z - (d - \frac{1}{2} c^T c) = 0. \end{cases}$$

- The null space variable y is eliminated from the constraint.
- We can solve y first with t, z fixed.

Space Reduction and Format Setting

$$\begin{cases} \min\limits_{t,y,z} & \frac{1}{2}t^2 + \frac{1}{2}y^T A_{uu} y + \frac{1}{2}z^T A_{vv} z + y^T A_{uv} z - f^T U y - f^T V z \\ \text{s.t.} & \frac{1}{2}z^T B_{vv} z - t - (B^T c)^T V z - (d - \frac{1}{2}c^T c) = 0. \end{cases}$$

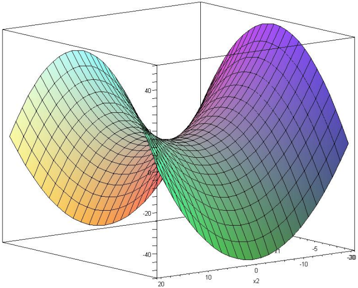
- The null space variable y is eliminated from the constraint.
- We can solve y first with t, z fixed.
- $A_{uu} \succeq 0$.
- If A_{uu} does not positive semi-definite, (QP1QC) is unbounded below.

Space Reduction and Format Setting

- For example, let
  $A = diag(-1, 0), B = diag(0, 1), c = (0, 0)^T, d = 0, f = (0, 0)^T.$
- $U = V = I$, $y = (x_1, 0)^T, z = (0, x_2)^T.$
- (QP1QC) formulated as

$$-\frac{1}{2}x_1^2 + \frac{1}{8}x_2^4$$

which the minimum value tends to $-\infty$.



Space Reduction and Format Setting

- The (QP1QC) in (t, z) becomes

$$\begin{cases} \min\limits_{t,z} & \frac{1}{2}t^2 + \frac{1}{2}z^T V^T \hat{A} V z - \hat{f}^T V z - \frac{1}{2} f^T U A_{uu}^+ U^T f \\ \text{s.t.} & \frac{1}{2} z^T B_{vv} z - t - (B^T c)^T V z - (d - \frac{1}{2} c^T c) = 0, \end{cases}$$

$\hat{A} = (I - AUA_{uu}^+ U^T)A$, $\hat{f} = (I - AUA_{uu}^+ U^T)f$, A_{uu}^+ is Moore-Penrose pseudoinverse of A_{uu}.

Space Reduction and Format Setting

- The (QP1QC) in (t, z) becomes

$$\begin{cases} \min\limits_{t,z} & \frac{1}{2}t^2 + \frac{1}{2}z^TV^T\hat{A}Vz - \hat{f}^TVz - \frac{1}{2}f^TUA_{uu}^+U^Tf \\ \text{s.t.} & \frac{1}{2}z^TB_{vv}z - t - (B^Tc)^TVz - (d - \frac{1}{2}c^Tc) = 0, \end{cases}$$

  $\hat{A} = (I - AUA_{uu}^+U^T)A$, $\hat{f} = (I - AUA_{uu}^+U^T)f$, A_{uu}^+ is Moore-Penrose pseudoinverse of A_{uu}.
- This is (QP1QC) in the lower dimensional case with $B_{vv} \succ 0$.

Space Reduction and Format Setting

- The (QP1QC) in (t, z) becomes

$$\begin{cases} \min\limits_{t,z} & \frac{1}{2}t^2 + \frac{1}{2}z^T V^T \hat{A} V z - \hat{f}^T V z - \frac{1}{2} f^T U A_{uu}^+ U^T f \\ \text{s.t.} & \frac{1}{2} z^T B_{vv} z - t - (B^T c)^T V z - (d - \frac{1}{2} c^T c) = 0, \end{cases}$$

  $\hat{A} = (I - AUA_{uu}^+ U^T)A$, $\hat{f} = (I - AUA_{uu}^+ U^T)f$, A_{uu}^+ is Moore-Penrose pseudoinverse of A_{uu}.

- This is (QP1QC) in the lower dimensional case with $B_{vv} \succ 0$.
- In this sense, we assume $B^T B \succ 0$ in (QP1QC).

Space Reduction and Format Setting

- By Cholesky factorization, there exists a nonsingular lower-triangular matrix P_1 such that $P_1^T(B^TB)P_1 = I$.
- $P_1^T A P_1 \in S_{n\times n}$
  $\Longrightarrow \exists P_2$ s.t. $P_2^T P_1^T A P_1 P_2 = diag(\alpha_1, \alpha_2, \cdots, \alpha_n)$,
  and $P_2^T P_1^T (B^TB) P_1 P_2 = I$.
- Let $P = P_1 P_2$, $P^TAP = diag(\alpha_1, \alpha_2, \cdots, \alpha_n), P^TB^TBP = I$.
- A and B^TB satisfies simultaneously diagonalized via congruence (SDC) if $B^TB \succ 0$. [9]

[9] By Feng et al., the quadratic programming with one quadratic constraint was solved when A and B^TB satisfies (SDC).

Space Reduction and Format Setting

- Let $w = P^{-1}x$, $\varphi = P^T f$, $\psi = P^T B^T c$, $\nu = d - \frac{1}{2}c^T c$.
- After (SDC), (DWP) can be written as the sum of separated squares

$$(P) \quad P_0 = \begin{array}{ll} \min & P(t,w) = \frac{1}{2}t^2 + \sum_{i=1}^{n}(\frac{1}{2}\alpha_i w_i^2 - \psi_i w_i) \\ \text{s.t.} & g(t,w) = 0, \end{array} \tag{4}$$

where $g(t,w) = \sum_{i=1}^{n}(\frac{1}{2}w_i^2 - \varphi_i w_i) - t - \nu$.

- We call (4) the primal problem (P).

Space Reduction and Format Setting

- The Lagrange function is

$$L(t, w; \sigma) = \frac{1}{2}t^2 + \sum_{i=1}^{n}\left(\frac{1}{2}(\alpha_i + \sigma)w_i^2 - (\psi_i + \sigma\varphi_i)w_i\right) - \sigma t - \sigma\nu,$$

where $\sigma \in R$ is the Lagrange multiplier.

Space Reduction and Format Setting

- The dual problem of (P) is formulated as

$$(D) \quad P_0^d = \quad \begin{array}{ll} \sup_\sigma & P^d(\sigma) = -\frac{1}{2}\sigma^2 - \frac{1}{2}\sum_{i=1}^n \frac{(\psi_i+\sigma\varphi_i)^2}{\alpha_i+\sigma} - \nu\sigma \\ \text{s.t.} & \sigma \in \mathcal{D} = \{\sigma \in R | \alpha_i + \sigma > 0\} = (\sigma_0, +\infty). \end{array} \tag{5}$$

- We call (5) the dual problem (D).
- $\sigma_0 = \max\{-\alpha_1, -\alpha_2, \cdots, -\alpha_n\}$.
- $P^d(\sigma)$ is a concave function.
- P_0^d may occurs as ① $\sigma^* \in \mathcal{D}$, ② $\sigma \to \sigma_0^+$, ③ $\sigma \to +\infty$.
- In case ③, this case never happen.

Analytic Solution to (DWP)

- The Lagrange function

$$L(t, w; \sigma) = \frac{1}{2}t^2 + \sum_{i=1}^{n}(\frac{1}{2}(\alpha_i + \sigma)w_i^2 - (\psi_i + \sigma\varphi_i)w_i) - \sigma t - \sigma\nu$$

  is minimize at $t(\sigma) = \sigma$ and $w(\sigma)_i = \frac{\psi_i + \sigma\varphi_i}{\alpha_i + \sigma}$.
- The point $(t(\sigma), w(\sigma), \sigma)$ is a saddle point such that $(t(\sigma), w(\sigma))$ and σ are optimal to (P) and (D), respectively if, and only if $g(t(\sigma), w(\sigma)) = 0$. [10]

[10] By the saddle point theorem in M.S. Bazaraa, H.D. Sherali, and C.M. Shetty, Nonlinear Programming Theory and Algorithms, 3rd. *Wiley Interscience*, 2006.

Analytic Solution to (DWP)

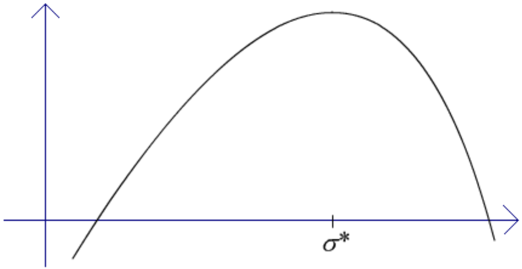
$$(D) \quad P_0^d = \sup_\sigma \quad P^d(\sigma) = -\frac{1}{2}\sigma^2 - \frac{1}{2}\sum_{i=1}^n \frac{(\psi_i+\sigma\varphi_i)^2}{\alpha_i+\sigma} - \nu\sigma$$
$$\text{s.t.} \quad \sigma \in \mathcal{D} = \{\sigma \in R | \alpha_i + \sigma > 0\} = (\sigma_0, +\infty).$$

- $\frac{dP^d(\sigma)}{d\sigma} = g(t(\sigma), w(\sigma))$.
- The point $(t(\sigma), w(\sigma))$ optimize the primal problem (P) if $\frac{dP^d(\sigma)}{d\sigma} = 0$.

Analytic Solution to (DWP)

In case ①,

- If $P_0^d = P^d(\sigma^*)$, it is necessary that $\frac{dP^d(\sigma^*)}{d\sigma} = 0$.
- $(t(\sigma^*), x(\sigma^*), \sigma^*)$ is a saddle point for the Lagrange function, where $t(\sigma^*) = \sigma^*$, $w(\sigma^*)_i = \frac{\psi_i + \sigma^* \varphi_i}{\alpha_i + \sigma^*}$.
- $x^* = Pw(\sigma^*)$ solves (DWP) with the optimal value $\frac{1}{2}(\sigma^*)^2 + \sum_{i=1}^n (\frac{1}{2}\alpha_i w(\sigma^*)_i^2 - \psi_i w(\sigma^*)_i)$.



Analytic Solution to (DWP)

In case ②, $P_0^d = \lim_{\sigma \to \sigma_0} P^d(\sigma)$.

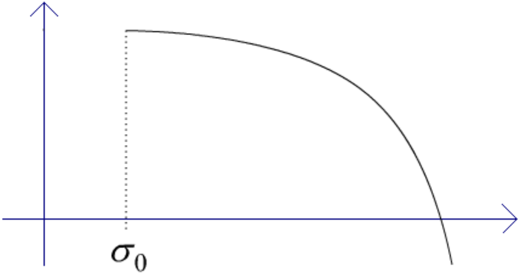
- $\{i | \alpha_i + \sigma_0 = 0\} \neq \emptyset$.
- Let $I := [1:k] = \{i | \alpha_i + \sigma_0 = 0\}$.
- $J := [k+1:n] = \{j | \alpha_j + \sigma_0 > 0\}$.
- Define $t(\sigma_0) = \sigma_0$ and

$$w(\sigma_0)_i = \lim_{\sigma \to \sigma_0^+} \frac{\psi_i + \sigma \varphi_i}{\alpha_i + \sigma} = \begin{cases} \varphi_i, & \text{if } i \in I, \\ \frac{\psi_i + \sigma_0 \varphi_i}{\alpha_i + \sigma_0}, & \text{if } i \in J. \end{cases}$$

- In this case, it may $\lim_{\sigma \to \sigma_0^+} \frac{dP^d(\sigma)}{d\sigma} = 0$ or $\lim_{\sigma \to \sigma_0^+} \frac{dP^d(\sigma)}{d\sigma} < 0$.

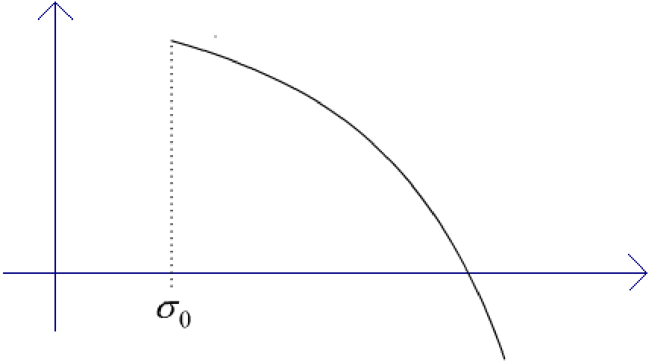
Analytic Solution to (DWP)

- If $P_0^d = \lim_{\sigma \to \sigma_0} P^d(\sigma)$ with $\lim_{\sigma \to \sigma_0^+} \frac{dP^d(\sigma)}{d\sigma} = 0$, then $(t(\sigma_0), x(\sigma_0), \sigma_0)$ is a saddle point for the Lagrange function.
- $x^* = Pw(\sigma_0)$ solves (DWP) with the optimal value $\frac{1}{2}(\sigma_0)^2 + \sum_{i=1}^{n}(\frac{1}{2}\alpha_i w(\sigma_0)_i^2 - \psi_i w(\sigma_0)_i)$.



Analytic Solution to (DWP)

- If $P_0^d = \lim_{\sigma \to \sigma_0} P^d(\sigma)$ with $\lim_{\sigma \to \sigma_0^+} \frac{dP^d(\sigma)}{d\sigma} < 0$, $(t(\sigma_0), w(\sigma_0))$ is not a feasible point of (P).



Analytic Solution to (DWP)

- We construct new point set by $(t(\sigma_0), w(\sigma_0))$:

$$S := \left\{ (\sigma_0, w^*) \middle| \begin{array}{l} w_j^* = \frac{\psi_j + \sigma_0 \varphi_j}{\alpha_j + \sigma_0},\ j \in J; \\ \sum_{i \in I} (\frac{1}{2}(w_i^*)^2 - \varphi_i w_i^*) = -\sum_{j \in J} (\frac{1}{2}(w_j^*)^2 - \varphi_j w_j^*) + \sigma_0 + \nu \end{array} \right\}.$$

- S forms the optimal solution of (P). [11]
- $x^* = Pw^*$ where $(t(\sigma_0), w^*) \in S$, is the optimal solution for (DWP).

[11] The though through the boundarification technique in Xing et al.

Dual of the Dual Problem

- Following we illustrate that the dual of the dual problem is a linearly constrained convex minimization problem (P^{dd}).

Dual of the Dual Problem

- Following we illustrate that the dual of the dual problem is a linearly constrained convex minimization problem (P^{dd}).
- (P^{dd}) is equivalent to (DWP) subject to additional n linear constraints.

Dual of the Dual Problem

- By rewrite the constraint of (D) as $\alpha_i + \sigma = r_i,\ \ i \in [1:n]$.
- The dual problem of (D) becomes

$$P_0^{dd} = \ -\sum_{i=1}^{n}(\psi_i - \alpha_i\varphi_i)\varphi_i + \left\{ \begin{array}{ll} \inf\limits_{\lambda\in R^n} & \sum\limits_{i=1}^{n}\alpha_i\lambda_i - \sum\limits_{i=1}^{n}|\psi_i - \alpha_i\varphi_i|\sqrt{2\lambda_i + \varphi_i^2} + \frac{1}{2}(\sum\limits_{i=1}^{n}\lambda_i - \nu)^2 \\ \text{s.t.} & \lambda_i + \frac{\varphi_i^2}{2} \geq 0,\ \forall i. \end{array} \right\}$$

- $\lambda_i \in R$ are the dual multipliers associated with the i^{th} linear equality constraint of (D).
- We call (P^{dd}) in short.

Dual of the Dual Problem

- By rewrite the constraint of (D) as $\alpha_i + \sigma = r_i, \;\; i \in [1:n]$.
- The dual problem of (D) becomes

$$P_0^{dd} = \;\; -\sum_{i=1}^{n}(\psi_i - \alpha_i\varphi_i)\varphi_i + \left\{ \begin{array}{ll} \inf\limits_{\lambda \in R^n} & \sum\limits_{i=1}^{n}\alpha_i\lambda_i - \sum\limits_{i=1}^{n}|\psi_i - \alpha_i\varphi_i|\sqrt{2\lambda_i + \varphi_i^2} + \frac{1}{2}(\sum\limits_{i=1}^{n}\lambda_i - \nu)^2 \\ \text{s.t.} & \lambda_i + \frac{\varphi_i^2}{2} \geq 0, \;\; \forall i. \end{array} \right\}$$

- $\lambda_i \in R$ are the dual multipliers associated with the i^{th} linear equality constraint of (D).
- We call (P^{dd}) in short.
- (P^{dd}) is a linearly constraint convex minimization problem.

Dual of the Dual Problem

- By the nonlinear transformation

$$w_i = \begin{cases} \varphi_i + \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i > 0 \\ \varphi_i - \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i < 0 \\ \varphi_i \pm \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i = 0, \end{cases} \tag{7}$$

(P^{dd}) is equivalent to (P) subject to n linearly constraint:

$$(\psi_i - \alpha_i\varphi_i)(w_i - \varphi_i) \geq 0.$$

Dual of the Dual Problem

- By the nonlinear transformation

$$w_i = \begin{cases} \varphi_i + \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i > 0 \\ \varphi_i - \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i < 0 \\ \varphi_i \pm \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i = 0, \end{cases} \quad (7)$$

  (P^{dd}) is equivalent to (P) subject to n linearly constraint:

$$(\psi_i - \alpha_i\varphi_i)(w_i - \varphi_i) \geq 0.$$

- In general, the nonlinear transformation is not one-to-one.

Dual of the Dual Problem

- By the nonlinear transformation

$$w_i = \begin{cases} \varphi_i + \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i > 0 \\ \varphi_i - \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i < 0 \\ \varphi_i \pm \sqrt{2\lambda_i + \varphi_i^2}, & \text{if } \psi_i - \alpha_i\varphi_i = 0, \end{cases} \tag{7}$$

  (P^{dd}) is equivalent to (P) subject to n linearly constraint:

$$(\psi_i - \alpha_i\varphi_i)(w_i - \varphi_i) \geq 0.$$

- In general, the nonlinear transformation is not one-to-one.
- This says that (P^{dd}) reveals the hidden convex nature of (P), which explains why the (DWP) has no duality gap.

Numerical Examples

Following we give some numerical examples to illustrate the solution of (DWP) by dual approach; the dual of the dual problem; and the relation between (DWP) and (P^{dd}).

Numerical Examples

Example 1 Let $A = -2, B = (0, -1)^T, c = (0, 2)^T, d = 14, f = 1$. (P) has the form:

$$min \quad P(w) = \frac{1}{2}(\frac{1}{2}w^2 + 2w - 12)^2 - w^2 - w$$

- The global minimum of (DWP) locates at $x^* = -7.748$ with the optimal value -49.109.

Numerical Examples

The dual problem (D) is

$$\begin{array}{ll} sup & P^d(\sigma) = -\frac{1}{2}\sigma^2 - \frac{(1-2\sigma)^2}{2\sigma-4} - 12\sigma \\ s.t. & \sigma \in \mathcal{D} = (2, \infty). \end{array}$$

- The supremum occurs at $\sigma^* = 2.522 \in \mathcal{D}$.
- The corresponding primal solution is $w(\sigma^*) = -7.748$.

Numerical Examples

The dual problem (D) is

$$\begin{array}{ll} sup & P^d(\sigma) = -\frac{1}{2}\sigma^2 - \frac{(1-2\sigma)^2}{2\sigma-4} - 12\sigma \\ s.t. & \sigma \in \mathcal{D} = (2, \infty). \end{array}$$

- The supremum occurs at $\sigma^* = 2.522 \in \mathcal{D}$.
- The corresponding primal solution is $w(\sigma^*) = -7.748$.

The dual of the dual problem is formulated as

$$\begin{array}{rl} P_0^{dd} = & -6 + \inf_\lambda[-2\lambda - 3\sqrt{2\lambda + 4} + \frac{1}{2}(\lambda - 12)^2] \\ s.t. & \lambda + 2 \geq 0, \end{array}$$

Numerical Examples

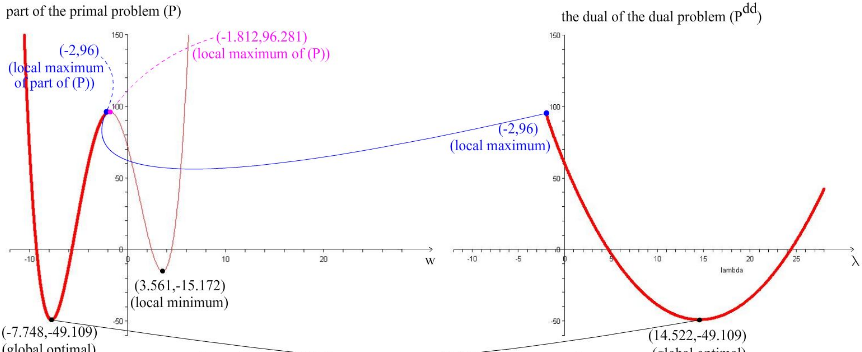


Figure 6 Figure 7

Numerical Examples

- The one-to-one nonlinear transformation is $w = -2 - \sqrt{2\lambda + 4}$, with we can rewrite (P^{dd}) as

$$\begin{array}{ll} \min & P(w) \\ \text{s.t.} & w + 2 \leq 0. \end{array}$$

Numerical Examples

- The one-to-one nonlinear transformation is $w = -2 - \sqrt{2\lambda + 4}$, with we can rewrite (P^{dd}) as

$$\begin{array}{ll} \min & P(w) \\ \text{s.t.} & w + 2 \leq 0. \end{array}$$

- This is the primal (P) subject to $w \leq -2$.
- The global minimum of part of (P) is mapped to the global minimum of (P^{dd}), which is also the global minimum of (P).

Numerical Examples

Example 2 Let $A = diag(1, -2), B = \begin{bmatrix} -0.07 & 0.04 \\ -0.01 & -1 \end{bmatrix}, c = (-2, 0)^T, d = 28, f = (-7, -22)^T$.

- After diagonalize A and $B^T B$ simultaneously via congruence, (P) has the form

$$\begin{aligned} min \quad P(w) = \; & \frac{1}{2}\left(\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - 1.998w_1 + 0.082w_2 - 26\right)^2 \\ & +101.035w_1^2 - 0.998w_2^2 + 98.285w_1 + 21.885w_2. \end{aligned}$$

Numerical Examples

The dual problem (D) is

$$\begin{array}{ll} sup & P^d(\sigma) = -\frac{1}{2}\sigma^2 - \frac{1}{2}[\frac{(-97.285+1.998\sigma)^2}{\sigma+202.071} + \frac{(-21.885-0.082\sigma)^2}{\sigma-1.997}] - 26\sigma \\ s.t. & \sigma \in \mathcal{D} = (1.997, \infty). \end{array} \quad (8)$$

- The supremum occurs at $\sigma^* = 4.8475 \in \mathcal{D}$.
- The corresponding primal solution $w(\sigma^*) = (-0.423, -7.817)^T$ is optimal to (P) with the optimal value -243.416.

Numerical Examples

The dual problem (D) is

$$\begin{array}{ll} sup & P^d(\sigma) = -\frac{1}{2}\sigma^2 - \frac{1}{2}[\frac{(-97.285+1.998\sigma)^2}{\sigma+202.071} + \frac{(-21.885-0.082\sigma)^2}{\sigma-1.997}] - 26\sigma \\ s.t. & \sigma \in \mathcal{D} = (1.997, \infty). \end{array} \quad (8)$$

- The supremum occurs at $\sigma^* = 4.8475 \in \mathcal{D}$.
- The corresponding primal solution $w(\sigma^*) = (-0.423, -7.817)^T$ is optimal to (P) with the optimal value -243.416.

The dual of the dual problem has the form

$$\begin{array}{rl} P_0^{dd} = & 999.529 + \inf_\lambda[202.071\lambda_1 - 1.997\lambda_2 - 501.088\sqrt{2\lambda_1 + 3.993} \\ & -22.049\sqrt{2\lambda_2 + 0.0067} + \frac{1}{2}(\lambda_1 + \lambda_2 - 26)^2] \\ s.t. & \lambda_1 \geq -1.9967,\ \lambda_2 \geq -0.0034. \end{array}$$

Numerical Examples

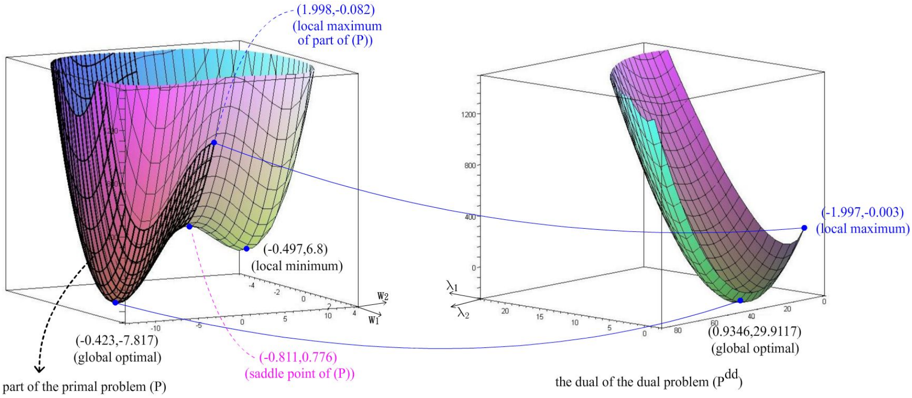


Figure 6

Figure 7

Numerical Examples

- The one-to-one nonlinear transformation is $w_1 = 1.998 - \sqrt{2\lambda_1 + 3.993}$ and $w_2 = -0.082 - \sqrt{2\lambda_2 + 0.0067}$, with we can rewrite (P^{dd}) as

$$\begin{array}{ll} \min & P(w) \\ \text{s.t.} & w_1 \leq 1.998, \;\; w_2 \leq -0.082. \end{array} \tag{9}$$

Numerical Examples

- The one-to-one nonlinear transformation is $w_1 = 1.998 - \sqrt{2\lambda_1 + 3.993}$ and $w_2 = -0.082 - \sqrt{2\lambda_2 + 0.0067}$, with we can rewrite (P^{dd}) as

$$\begin{array}{ll} \min & P(w) \\ \text{s.t.} & w_1 \leq 1.998, \; w_2 \leq -0.082. \end{array} \tag{9}$$

- This is the primal (P) subject to $w_1 \leq 1.998, \; w_2 \leq -0.082$.
- The graph of (9) is superimposed with the black bold net.
- The nonlinear transformation sends (P^{dd}) back to part of (P).
- The global minimum of part of (P) is mapped to the global minimum of (P^{dd}), which is also the global minimum of (P).

Numerical Examples

Example 3 (The Mexican hat) Let
$A = 0_{2\times 2}, B = diag(-0.5, -0.5), c = (0,0)^T, d = 38, f = (0,0)^T$. The primal problem (P) is

$$min \quad P(w) = \frac{1}{2}(\frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 - 38)^2.$$

- The global minimum is on the circle $\{(w_1, w_2) | \frac{1}{2}w_1^2 + \frac{1}{2}w_2^2 = 38\}$ with the optimal value 0.
- There is a local maximum at $(0, 0)$.

Numerical Examples

The dual problem (D) is

$$\begin{array}{ll} sup & P^d(\sigma) = -\frac{1}{2}\sigma^2 - 38\sigma \\ s.t. & \sigma \in \mathcal{D} = (0, \infty). \end{array}$$

- The supremum occurs at $\sigma_0 = 0$, with $\lim_{\sigma \to 0^+} \frac{d\mathrm{P}^d(\sigma)}{d\sigma} = -38 < 0$.
- By constructing the solution set as in (2) (boundarification technique), we get the same optimal solution set as (DWP) $\{(w_1, w_2) | \frac{1}{2}(w_1)^2 + \frac{1}{2}(w_2)^2 = 38\}$, with the optimal value 0.

Numerical Examples

The dual problem (D) is

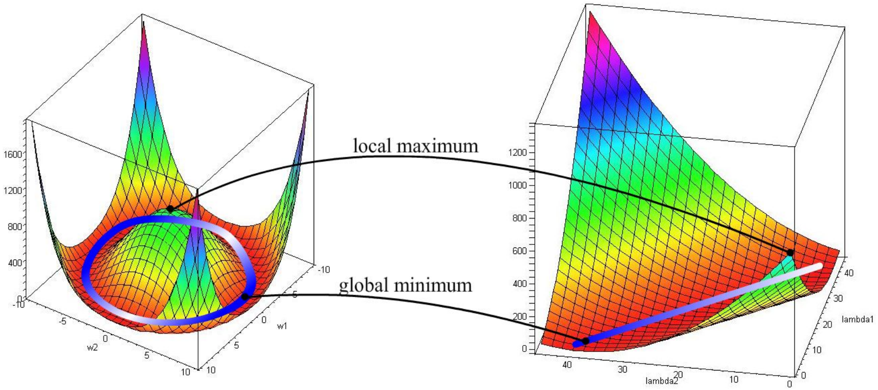
$$\begin{array}{ll} sup & P^d(\sigma) = -\frac{1}{2}\sigma^2 - 38\sigma \\ s.t. & \sigma \in \mathcal{D} = (0, \infty). \end{array}$$

- The supremum occurs at $\sigma_0 = 0$, with $\lim_{\sigma \to 0^+} \frac{d\mathrm{P}^d(\sigma)}{d\sigma} = -38 < 0$.
- By constructing the solution set as in (2) (boundarification technique), we get the same optimal solution set as (DWP) $\{(w_1, w_2) | \frac{1}{2}(w_1)^2 + \frac{1}{2}(w_2)^2 = 38\}$, with the optimal value 0.

The problem P^{dd} has the form

$$\begin{array}{rl} P_0^{dd} = & \inf_\lambda \frac{1}{2}(\lambda_1 + \lambda_2 - 38)^2 \\ s.t. & \lambda_1 \geq 0, \ \lambda_2 \geq 0. \end{array}$$

Numerical Examples



The primal problem (P)	The dual of the dual problem (P^{dd})
Figure 8	Figure 9

Numerical Examples

- The nonlinear transformation is $w_i = \pm\sqrt{2\lambda_i}, i = 1, 2.$
- It is not a one-to-one transformation.

Numerical Examples

- The nonlinear transformation is $w_i = \pm\sqrt{2\lambda_i}, i = 1, 2$.
- It is not a one-to-one transformation.
- It maps (P^{dd}) back to the entire primal problem with no additional constraint.
- The optimal solution set of (P^{dd}) is collapsed into the line segment $\{(\lambda_1, \lambda_2) | \lambda_1 + \lambda_2 = 38, \ \lambda_1 \geq 0, \ \lambda_2 \geq 0\}$.

Thank you for listening!
Any question?