6th European Conference on Computational Mechanics (ECCM 6)
7th European Conference on Computational Fluid Dynamics (ECFD 7)
1115 June 2018, Glasgow, UK

A MODIFIED SEARCH DIRECTION METHOD WITH WEAKLY IMPOSED KARUSH-KUHN-TUCKER CONDITIONS FOR GRADIENT BASED CONSTRAINT OPTIMIZATION FOR VERY LARGE PROBLEMS

Long Chen[1], Armin Geiser[1], Roland Wüchner[1] and Kai-Uwe Bletzinger[1]

[1] Chair of Structural Analysis, Technical University of Munich
Arcisstr. 21, D-80333 München
long.chen@tum.de, https://www.st.bgu.tum.de/

Key words: Gradient-based constrained optimization, Karush-Kuhn-Tucker conditions, singular-value decomposition, shape optimization, Vertex Morphing

Abstract. Motivated by the applications of the Vertex Morphing method [1] for very large shape optimization problems, where various response evaluations and multiple physics are often considered in a constrained optimization, we propose a robust modified search direction method for the optimization procedure. The solution of a general constrained gradient-based optimization problem satisfies the necessary Karush-Kuhn-Tucker (KKT) conditions [2]. There exist numerous methods to solve constrained optimization problems, which try to travel along the active constraint to find the local minimum [2][3]. This might lead to inefficiency in the optimization process for very large problems. In the proposed method, the KKT conditions are weakly imposed in each optimization step. The search direction is modified and designed to find a solution where the KKT conditions can be better fulfilled compared using the steepest descent direction. To accomplish this, the singular-value decomposition method [4] is applied to both the objective and constraint sensitivity. The results are shown first with analytical 2D problems and then the results of shape optimization problems with a large number of design variables are discussed. In order to robustly deal with complex geometries, the Vertex Morphing method is used.

1 INTRODUCTION

In the practice of shape optimization using the Vertex Morphing method [1], a large number of design variables is often considered in a constrained optimization. The optimization problem as well as the constraints can be highly non-linear. Various methods can be used for constrained optimization, such as the feasible direction method, gradient projection method, Sequential Linear Programming (SLP) and Sequential Quadratic Programming (SQP). When treating inequality constraints, the active-set strategy is widely used in the above mentioned method, since it can conveniently generalize the equality constraint to an inequality constraint [3]. The optimization algorithms utilize the steepest descent direction $\mathbf{c} = -\Delta f^T(\mathbf{x})$ of the unconstrained problem $f(\mathbf{x})$ as the search direction if there is no active constraint. When the constraints are active, their effect are then included in calculating the search direction. The algorithms try to find a local minimum by traveling along the active constraint. In the context of shape optimization with a large number of design variables, due to the high nonlinearity of the optimization problem as well as a high requirement regarding the mesh quality, the existing optimization algorithms might result in inefficiency in the optimization procedure. In the present work, we propose a modified search direction method: instead of using the steepest-descent direction when there is no active constraint, we consider the impact of the constraint. The search direction is modified so that the Karush-Kuhn-Tucker (KKT) gradient conditions can be better fulfilled in the next iteration compared to the steepest-descent direction. This is accomplished by using the singular-value decomposition method for the objective and constraint gradients. The proposed method is demonstrated using 2D analytical constrained optimization examples. Moreover, selected results of shape optimizations of shells with Vertex Morphing method with various geometrical constraints are discussed.

2 Karush-Kuhn-Tucker optimality conditions

The necessary conditions for an equality- and inequality-constrained problem can be summed up in what are commonly known as the *Karush-Kuhn-Tucker (KKT) optimality conditions*. The KKT conditions are repeated as follows for the sake of clarity [2]:

Let $\mathbf{x}^\star$ be a regular point of the feasible set that is a local minimum for $f(\mathbf{x})$, subject to $h_i(\mathbf{x}) = 0$; $i = 1$ to p; $g_j(\mathbf{x}) \leq 0$; $j = 1$ to m. Then there exist Lagrange multipliers $\boldsymbol{\lambda}^\star$ (a p-vector) and $\boldsymbol{\mu}^\star$ (an m-vector) such that the Lagrangian function is stationary with respect to x_k, λ_i, and μ_j at the point

1. Lagrangian function for the problem written in the standard form:

$$\begin{aligned} L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) =& f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i h_i(\mathbf{x}) + \sum_{j=1}^{m} \mu_j g_j(\mathbf{x}) \\ =& f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathbf{T}} \mathbf{h}(\mathbf{x}) + \boldsymbol{\mu}^{\mathbf{T}} \mathbf{g}(\mathbf{x}) \end{aligned} \tag{1}$$

2. Gradient conditions:

$$\frac{\partial L}{\partial x_k} = \frac{\partial f}{\partial x_k} + \sum_{i=1}^{p} \lambda_i^* \frac{\partial h_i}{\partial x_k} + \sum_{j=1}^{m} \mu_j^* \frac{\partial g_j}{\partial x_k} = 0;\ \mathrm{k = 1\ to\ n} \tag{2}$$

$$\frac{\partial L}{\partial \lambda_i} = 0 \Rightarrow h_i(\mathbf{x}^*) = 0;\ \mathrm{i = 1\ to\ p} \tag{3}$$

$$\frac{\partial L}{\partial \mu_j} = 0 \Rightarrow g_j(\mathbf{x}^*) = 0;\ \mathrm{j = 1\ to\ m} \tag{4}$$

3. Feasibility check for inequalities:

$$h_i(\mathbf{x}^*) = 0;\ \mathrm{i = 1\ to\ p};\ g_j(\mathbf{x}^*) \leq 0;\ \mathrm{j = 1\ to\ m} \tag{5}$$

4. Switching conditions:

$$\mu_j^* g_j(\mathbf{x}^*) = 0;\ \text{j} = 1 \text{ to m} \tag{6}$$

5. Non-negativity of Lagrange multipliers for inequalities:

$$\mu_j^* \geq 0;\ \text{j} = 1 \text{ to m} \tag{7}$$

6. Regularity check: Gradients of the active constraints must be linearly independent. In such a case the Lagrange multipliers for the constraints are unique.

3 Singular-value decomposition

In the proposed method, the KKT optimality conditions are weakly imposed in the determination of the search direction at each iteration. To accomplish this, the singular-value decomposition (SVD) method is used. Therefore, a short review on the SVD method is given in this section.

3.1 Basics on singular-value decomposition

The singular-value decomposition is the generalization of the eigendecomposition of a positive semidefinite normal matrix to any $m \times n$ matrix via an extension of the polar decomposition [4]. Formally, the singular-value decomposition of an $m \times n$ real or complex matrix $\mathbf{M}$ is a factorization of the form $\mathbf{U\Sigma V}^*$,

$$\mathbf{M} = \mathbf{U\Sigma V}^* \tag{8}$$

where $\mathbf{M}$ is a $m \times n$ real or complex matrix, $\mathbf{U}$ is a $m \times m$ real or complex unitary matrix, $\mathbf{\Sigma}$ is a $m \times n$ rectangular diagonal matrix with non-negative real numbers on the diagonal, $\mathbf{V}$ is a $n \times n$ real or complex unitary matrix and $\mathbf{V}^*$ is the conjugated transpose of $\mathbf{V}$.

The diagonal entries σ_i of $\mathbf{\Sigma}$ are singular values of $\mathbf{M}$. The columns of $\mathbf{U}$ and the columns of $\mathbf{V}$ are called the *left-singular vectors* and *right-singular vectors* of $\mathbf{M}$, respectively.

Since $\mathbf{U}$ and $\mathbf{V}^*$ are unitary, the columns of each of them form a set of orthonormal vectors, which can be regarded as basis vectors. The matrix $\mathbf{M}$ maps the basis vector $\mathbf{V}_i$ to the stretched unit vector $\sigma_i \mathbf{U}_i$. By the definition of a unitary matrix, the same is true for their conjugated transposes $\mathbf{U}^*$ and $\mathbf{V}$, except the geometric interpretation of the singular values as stretches is lost. In short, the columns of $\mathbf{U}$, $\mathbf{U}^*$, $\mathbf{V}$, $\mathbf{V}^*$, are orthonormal basis.

3.2 Use of SVD to analyze input-output systems

Any real matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ defines an input-output system as follows

$$\mathbf{z} = \mathbf{Ay} \tag{9}$$

with $\mathbf{z} \in \mathbb{R}^n$ denoting the output vector and $\mathbf{y} \in \mathbb{R}^m$ denoting the input vector. Singular-value decomposition of the matrix $\mathbf{A}$ can be used to analyze the input-output system,

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \sum_{i=1}^{min(m,n)} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{10}$$

Assume that an input vector corresponds to the $k-th$ right singular vector, which is scaled by a factor d with $\mathbf{y} = d\mathbf{v}_k$. Due to the fact that $\mathbf{u}_i$ and $\mathbf{v}_i$ are orthonormal bases we can obtain the output vector as follows [6],

$$\mathbf{z} = \left(\sum_{i=1}^{min(m,n)} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \right) \mathbf{y}_k = d\sigma_{kk}\mathbf{u}_k \tag{11}$$

with $\|\mathbf{z}\|_2 = \|d\sigma_{kk}\mathbf{u}_k\|_2 = d\sigma_{kk}$.

4 The SVD modified search direction method

In the present method, we propose a new way for finding the search path towards the local minimum. Instead of using the steepest descent direction of the unconstrained problem in the feasible domain (i.e. there is no active constraint), we take the inactive constraint into account and use it to modify the search direction. The goal is to find a searching path, which can possibly avoid the part of traveling along the active constraint and find the local minimum in a robust way.

4.1 Singular-value decomposed objective and constraint sensitivity

The gradients of the objective and the constraint can be obtained by sensitivity analysis. Let the row vector $\mathbf{c}$ and $\mathbf{a}$ represent the gradients of the objective and inequality constraint, respectively. They are given as:

$$\mathbf{c} = \nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & ... & \frac{\partial f}{\partial x_n} \end{pmatrix} \tag{12}$$

$$\mathbf{a} = \nabla g = \begin{pmatrix} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & ... & \frac{\partial g}{\partial x_n} \end{pmatrix} \tag{13}$$

We construct the sensitivity matrix $\mathbf{m}$ from the normalized objective sensitivity $\hat{\mathbf{c}}$ and the normalized constraint sensitivity $\hat{\mathbf{a}}$:

$$\mathbf{m} = \begin{pmatrix} \hat{\mathbf{c}} \\ \hat{\mathbf{a}} \end{pmatrix} \tag{14}$$

We apply singular-value decomposition as shown in equation (8) to the sensitivity matrix $\mathbf{m}$,

$$\mathbf{m} = \mathbf{u}\mathbf{s}\mathbf{v}^* \tag{15}$$

where matrix $\mathbf{u}$ contains the left-singular column vectors and $\mathbf{v}$ contains the right-singular vectors of the sensitivity matrix $\mathbf{m}$. We write the $i-th$ left-singular column vector $\mathbf{u}_i$, the $i-th$ singular value s_i, and the $i-th$ right-singular column vector $\mathbf{v}_i$ in a set m_i

$$m_i = \{\mathbf{u}_i, s_i, \mathbf{v}_i\} \tag{16}$$

We denote the set m_i as the $i-th$ *design mode* of the sensitivity matrix $\mathbf{m}$.

4.2 A 2D analytical example demonstrating the SVD modified search direction method

A 2D analytical optimization example is introduced and will be used to demonstrate the proposed method.

Consider the following analytical optimization problem,

Minimize

$$(x_1 - 2)^2 + (x_2 - 2)^2, \tag{17}$$

which is subject to the constraint

$$-\frac{1}{10}(x_1 - 3)^2 - x_2 + 3 \leq 0. \tag{18}$$

4.2.1 Singular-value decomposition of the sensitivity matrix m

We choose an initial design $\mathbf{x}_0 = (-12, -4)$ and its sensitivity matrix can be calculated

$$\mathbf{m}_0 = \begin{pmatrix} -0.9191 & -0.3939 \\ 0.9487 & -0.3162 \end{pmatrix} \tag{19}$$

Applying SVD to the sensitivity matrix $\mathbf{m}_0$ we obtain the right-singular vectors $\mathbf{v}_1$ and $\mathbf{v}_2$,

$$\mathbf{v}_1 = \begin{pmatrix} 0.9991 \\ 0.0416 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} -0.0416 \\ 0.9991 \end{pmatrix} \tag{20}$$

the left-singular vectors $\mathbf{u}_1$ and $\mathbf{u}_2$,

$$\mathbf{u}_1 = \begin{pmatrix} -0.7071 \\ 0.7071 \end{pmatrix} \quad \mathbf{u}_2 = \begin{pmatrix} -0.7071 \\ -0.7071 \end{pmatrix} \tag{21}$$

as well as the singular values,

$$s_1 = 1.3219 \quad s_2 = 0.5026 \tag{22}$$

Recalling the properties discussed in the section 3.2, we can interpret the two design modes as follows:

- the first design mode $m_1 = \{\mathbf{u}_1, s_1, \mathbf{v}_1\}$: by taking $\delta\mathbf{v}_1$ as design change, we can obtain a change in objective as well as in constraint function $\delta\mathbf{J} = s_1\delta\mathbf{u}_1$, which is a decrease in the objective function and an increase in the constraint function. We denote the first mode as *the primal mode*.

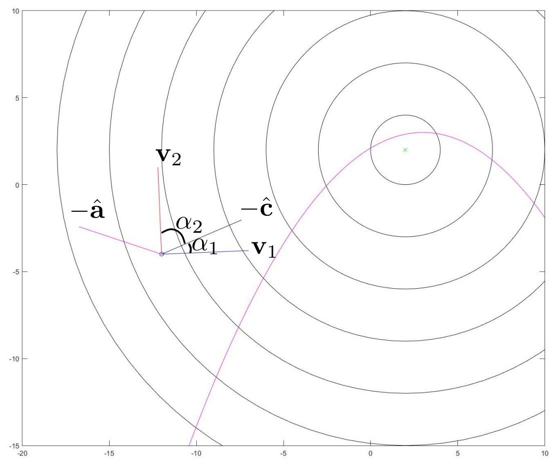

Figure 1: 2D analytical optimization with the design modes

- the second design mode $m_2 = \{\mathbf{u}_2, s_2, \mathbf{v}_2\}$: by taking $\delta\mathbf{v}_2$ as design change, we can obtain a change in objective as well as in constraint function $\delta\mathbf{J} = s_2\delta\mathbf{u}_2$, which is a decrease in both the objective and the constraint function. We denote the second mode as *the dual mode*.

The two design changes $\mathbf{v}_1$ and $\mathbf{v}_2$ from the design modes are illustrated in figure 1. The circles in figure 1 indicate the contour lines of the objective function described in equation (17) and the parabola indicates the constraint described in equation (18).

Recall that at the local optimum, the design satisfies the necessary KKT condition as described in the section 2. We are especially interested in the gradient conditions described in equation (2). Rewriting the KKT gradient condition for our analytical example we get

$$\mathbf{c} = \mu\mathbf{a} \tag{23}$$

where μ is the Lagrange Multiplier.

In this case, by applying SVD to the sensitivity matrix we obtain $\mathbf{v}_1 = \alpha\hat{\mathbf{a}}$ with $\alpha \in \mathbb{R}$ and $\mathbf{v}_2 = \mathbf{0}$ since the normalized objective gradient and normalized constraint gradient are parallel.

4.2.2 Modified search direction using the design modes

We denote the angle between $\mathbf{v}_1$ and $-\hat{\mathbf{c}}$ as α_1 and the angle between $\mathbf{v}_2$ and $-\hat{\mathbf{c}}$ as α_2. Thus we can rewrite the normalized steepest descent direction $-\hat{\mathbf{c}}$ as

$$-\hat{\mathbf{c}} = \cos\alpha_1\mathbf{v}_1 + \cos\alpha_2\mathbf{v}_2 \tag{24}$$

In order to get a design which better fulfills the KKT gradient condition in the next iteration compared to using the steepest descent direction, we modify the search direction as

$$\mathbf{s} = \cos\alpha_1\mathbf{v}_1 + c\cos\alpha_2\mathbf{v}_2 \tag{25}$$

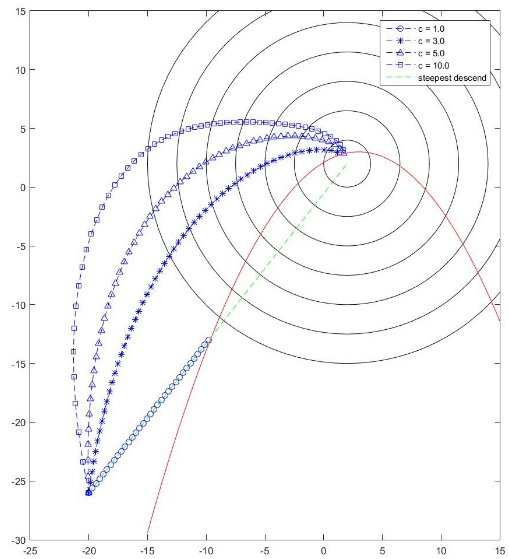

Figure 2: 2D analytical optimization with different parameter c

where the factor $c > 1$ is introduced to enlarge the contribution of the *dual mode* design change $\mathbf{v}_2$. We define the factor c as *duality factor* and its influence on the optimization procedure is demonstrated in the following subsection.

4.2.3 Results of the 2D analytical example

Applying the SVD modified search direction method we obtain the results of the 2D analytical optimization problem illustrated in subsection 4.2. In figure 2 we show the different optimization progresses for different duality factors c. It can be obviously observed, the bigger the duality factor c is chosen, the bigger is the impact of the constraint gradient on the modified search direction. It should be noted, however, that it is not guaranteed to achieve the local minimum with every duality factor $c > 1$.

Figure 3 shows the optimization progresses with different initial designs. The duality factor $c = 5.0$ is chosen. Starting from the six different initial designs, the SVD modified search direction method can successfully achieve the local minimum.

5 Vertex Morphing method for shape optimization

In the present work, the SVD modified search direction method is applied in node based shape optimization with a large number of design variables. In order to robustly deal with complex geometry, the Vertex Morphing method [1][5] is used. It is shortly reviewed in the following.

The idea of Vertex Morphing is to control the surface nodes $\mathbf{x} = [x_1, x_2, ..., x_n]$ with a design control field $\mathbf{s} = [s_1, s_2, ..., s_n]$, filtered with a filter function $A(x, x_0, r)$. The filter radius r acts as a design handle and controls the smoothness of the resulting surface. The so called forward mapping step using the linear mapping matrix $\mathbf{A}$ is defined as follows,

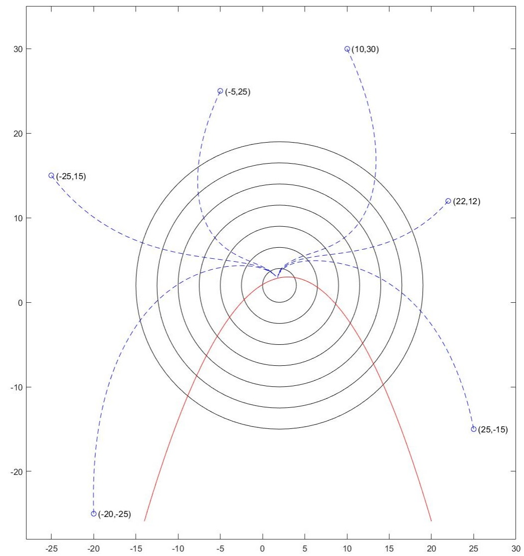

Figure 3: SVD modified search direction optimization with the duality factor $c = 5$ applied for different initial designs

$$x_i = A_{ij} s_j \tag{26}$$

Similarly, the change of the control field $\delta\mathbf{s}$ is mapped on to the change of the design surface $\delta\mathbf{x}$

$$\delta x_i = A_{ij} \delta s_j \tag{27}$$

The nodes of the control field are the design variables of the gradient based optimization. Following the chain rule of differentiation, the sensitivities of the response J w.r.t. the geometry $\mathbf{x}$ are backward mapped to the design control field using the adjoint or backward mapping matrix $\mathbf{A}^*$, with $\mathbf{A}^* = \mathbf{A}^T$ for regular grids.

$$\frac{dJ}{ds_i} = \frac{dJ}{dx_j}\frac{dx_j}{ds_i} = A_{ji}\frac{dJ}{dx_j} \tag{28}$$

Using equations (27) and (28) the design surface is modified iteratively, ensuring smooth shape updates.

6 Numerical example of large shape optimization problem

The presented method has been tested through several constrained shape optimization problems. The example shown is a 3D shell structure under static loading illustrated in figure 4. The cylinder shell with a radius $R = 10$ and a thickness $t = 1$ is supported with a fixed support at the left bottom edge and a roller support at the right. A single point load $p = 1e5$ is applied at the center of the shell structure. The Young's Modulus $E = 2.068e11$ and the Poisson's ratio $\mu = 0.29$ are assigned to the shell material property.

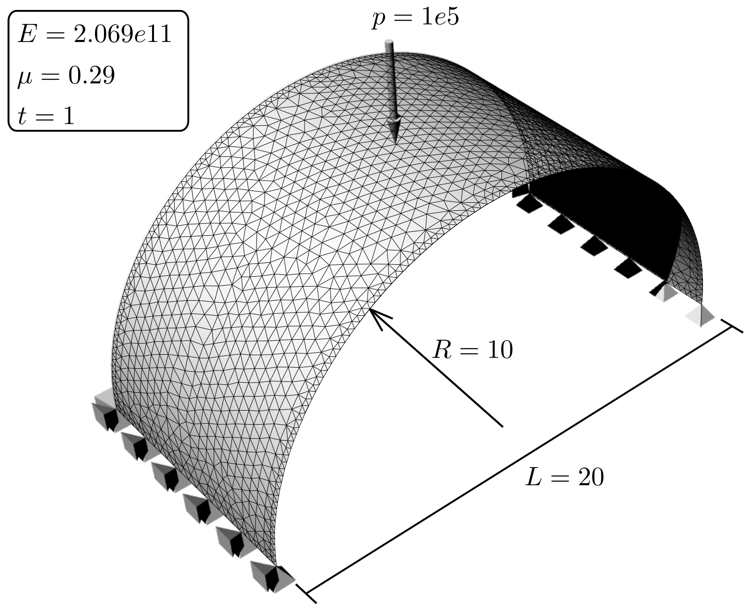

Figure 4: 3D shell structure unconstrained strain energy optimization

A linear elastic static analysis is carried out and a co-rotational 3-node shell element based on the Assumed Natural Deviatoric Strain formulation is used. The strain energy is set to be the objective function, while different geometrical constraints are considered. The shape of the structure is optimized and the Vertex Morphing method is used.

6.1 Unconstrained strain energy optimization for 3D shell structure

Carrying out the unconstrained optimization with the Vertex Morphing technology, where the steepest-descent direction of strain energy objective function with a fixed step size is used as design update, a result characterized by a filter radius r can be obtained. This is illustrated in figure 5 on the right side. After 61 steps the optimization is converged to a local minimum and the strain energy is decreased to 1.99% of the initial design.

6.2 Constrained strain energy optimization for 3D shell structure with one geometrical constraint

We add a geometric constraint to the unconstrained problem discussed in the previous subsection 6.1. The geometric constraint is a sphere as is shown in figure 6. In order to apply the SVD modified search direction method, we add an artificial gradient for the geometric constraint,

$$\frac{\partial g}{\partial x_i} = \frac{\partial d^2}{\partial x_i}(\frac{b-d}{b})^p \tag{29}$$

where d is the distance function, b is the predefined distance in which the geometric constraint gradient is calculated ($d \leq b$), and $p \in \mathbb{R}$ is the exponential factor which

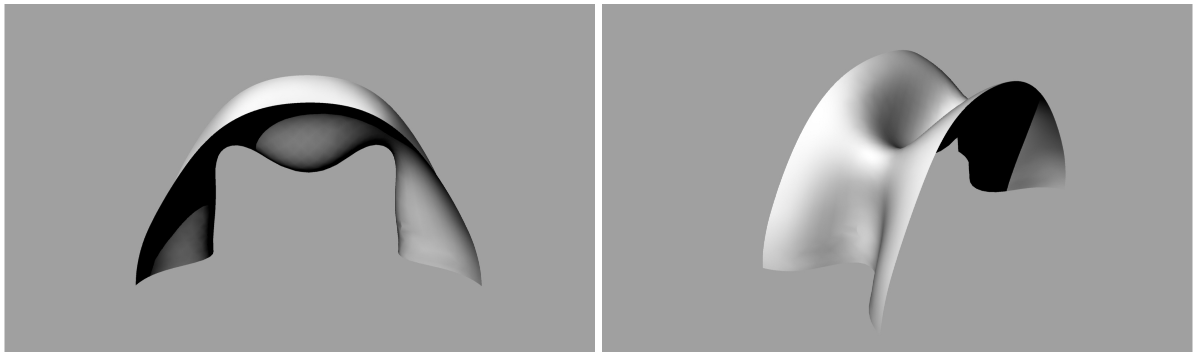

Figure 5: 3D shell structure unconstrained strain energy optimization

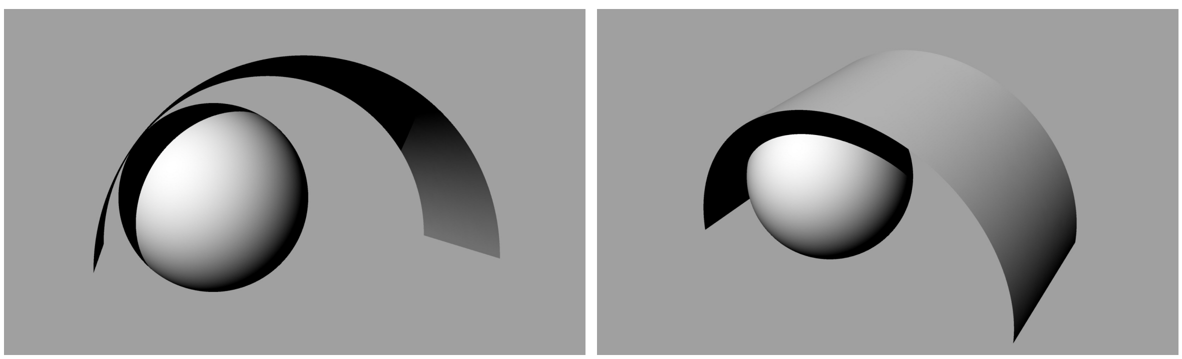

Figure 6: 3D shell structure initial design with spheric geometry constraint

penalizes the gradient.

The optimization converged after 66 steps and the strain energy is reduced to 4.04% of the initial design. The optimized geometry is shown in figure 7. The objective function value for the initial design is at a local maximum, due to the flat shape of the shell structure. By weakly imposing the KKT gradient conditions, the algorithm is aware of the incoming constraint and be able to take another path before the constraint gets active.

The optimization problem and the constraint of this example is highly non-linear. By taking the constraint information into account a priori, the whole optimization progress is regularized and is therefore smooth. As can be seen in the optimization result, the shape of the whole structure morphs according to the incoming constraint due to the regularization of the optimization problem. A sudden change in design update due to the KKT *switching conditions* as described in equation (6) is therefore avoided.

6.3 Constrained strain energy optimization for 3D shell structure with two geometrical constraints

We add a second sphere geometric constraint to the constrained optimization problem discussed in the previous subsection as is shown in figure 8. The optimization converged

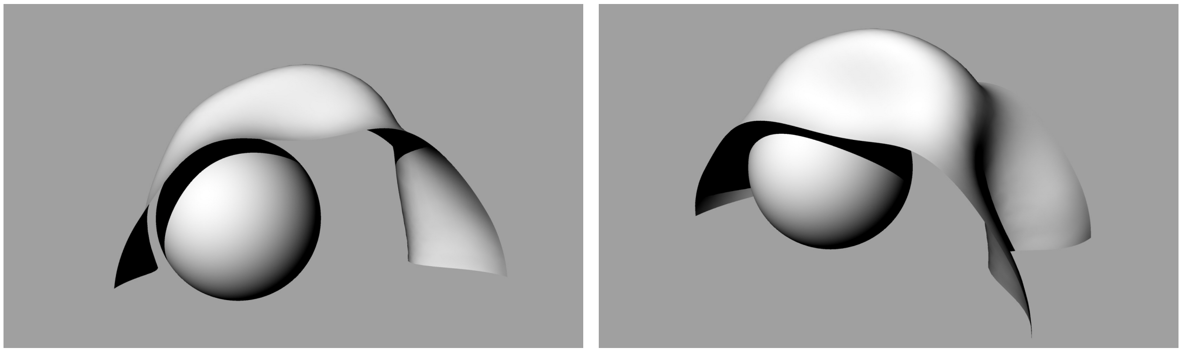

Figure 7: Optimized design for the 3D shell structure with spheric geometry constraint

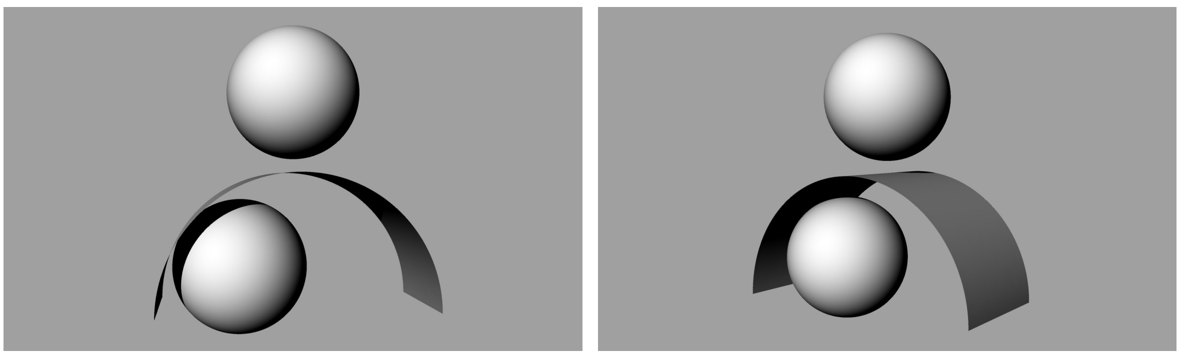

Figure 8: 3D shell structure initial design with two spheric geometry constraint

after 49 steps and the strain energy is reduced to 2.56% of the initial design. The structure morphs towards a local minimum where a violation of the two geometrical constraints is effectively avoided. Again, one identifies the regularization of the constrained optimization problem, the overall shape morphs smoothly with the goal reducing the objective value and keeping the constraints from being violated. Compared to the example with single geometrical constraint, a better local minimum is obtained due to the high nonlinearity of the optimization problem.

7 Conclusions

In the present work, a modified search direction method for constrained optimization is developed. We construct the modified search direction in feasible domain by utilizing the *primal* and *dual mode* computed by the singular-value decomposition of the sensitivity matrix. The *duality factor* is introduced in order to enlarge the contribution of the *dual mode*, such that the KKT gradient conditions can be better fulfilled at the next iteration compared to the direct use of the steepest descent search direction. Various constrained optimization examples show that the proposed method finds local minimums by properly avoiding the constraint violations.

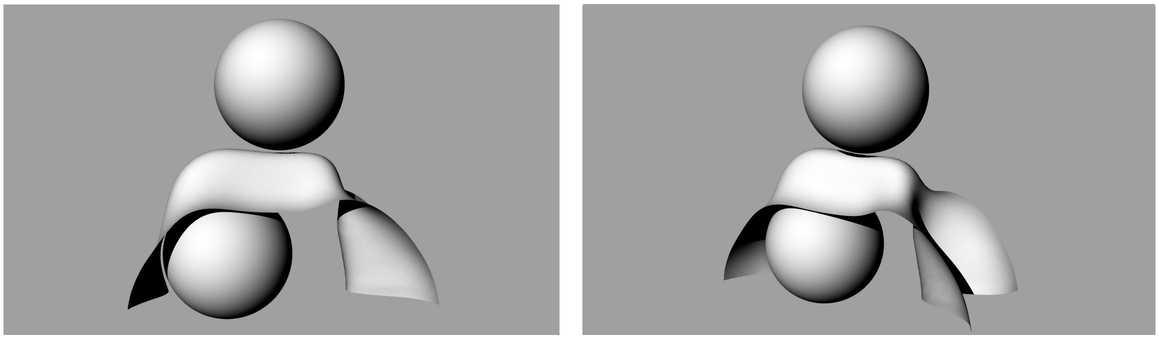

Figure 9: Optimized design for the 3D shell structure with two spheric geometry constraint

The proposed algorithm regularizes the constrained optimization problem and provides smooth design updates. A discontinuity in the search direction resulting from the KKT *switching condition* compared to active-set strategy can be effectively avoided. It is shown in the results that the proposed method is very promising when dealing with highly non-linear, large optimization problems.

REFERENCES

[1] Bletzinger, K.-U., A consistent frame for sensitivity filtering and the vertex assigned morphing of optimal shape. *Struct. Multidiscip. Optim.* (2014) **49**:873–895.

[2] Arora, J. S. *Introduction to Optimum Design*. Academic Press, 3rd Edition, 2011.

[3] Gallagher, R.H. and Zienkiewicz, O.C. *Optimum Structural Design*. John Wiley & Sons, 1973.

[4] Banerjee, S. and Roy, A. *Linear Algebra and Matrix Analysis for Statistics*. CRC Press, 2014.

[5] Hojjat, M., Stavropoulou, E. and Bletzinger, K-U., The Vertex Morphing method for node-based shape optimization. *Comput. Methods Appl. Mech. Eng.* (2014) **268**: 494–513.

[6] Gerzen N. *Analysis and Applications of Variational Sensitivity Information in Structural Optimisation*. PhD Thesis, Dortmund University of Technology, 2014.